Agneta Gulz

The Planning of Action as a Cognitive and Biological Phenomenon

Lund University Cognitive Studies; 2
1991

AGNETA GULZ
Cognitive Science Dept., Lund University, Sweden

THE PLANNING OF ACTION AS A COGNITIVE AND BIOLOGICAL PHENOMENON

KOGNITIONSFORSKNING, LUNDS UNIVERSITET
KUNGSHUSET/LUNDAGÅRD
S-223 50 LUND, SWEDEN

Illustrations: Magnus Haake
Figures: Christian Balkenius

Printed in Sweden
ISSN 1101-8453, ISBN 91-628-0224-0

For Tobbe,
with whom I shared plans and dreams for a long time. Thank you for the experiences we had and for what we learned together.

For Lennart,
who was the first person to educate me and help me to understand that one's own dreams and plans are valuable and that they can be realized.

TABLE OF CONTENTS

PREFACE

A main origin of this thesis was simply my interest in and fascination with the phenomenon of human planning and my feeling that the planning of action is an important feature in the life of many people. Another important background factor was my interest in comparative psychology and in studying psychological phenomena from an evolutionary perspective. I believe that important aspects of the cognitive and behavioural capacities in humans are products of evolution and may fruitfully be studied by being put into that context.

My main early influence came from the writings of Konrad Lorenz. Later I profited from a stay at the Department of Psychology, biomathematical section, at Ztirich University, during 1986-1987, where I became further acquainted with this tradition. There I was more thoroughly introduced both to the field of comparative psychology, developmental and evolutionary, and to the field of motivational psychology. For this influence, I would first of all like to thank Norbert Bischof and Doris Kohler-Bischof.

During the spring and summer of 1988, I had the opportunity to study at the School of Cognitive Science at Sussex University. I would, in particular, like to thank Andy Clark, David Pickles and Aaron Sloman for inspiring and useful ideas and discussions.

During the last couple of years, I have been working on this dissertation at the Department of Cognitive Science at the University of Lund- one of the most enjoyable and inspiring research environments that I can imagine. Thank you all for being there and for being who you are. Specifically, I want to thank Christian Balkenius for stimulating discussions, for understanding so much about the Berry-Creatures and for all his work in constructing the figures and interpreting my scribbles; Kenneth Holmquist for his encouragement and patience and for scanning in the illustrations; Linus Brostrom for invaluable support of all kinds and for his thoroughness and sharpness in reading drafts; Lukas Book for inspiring and valuable discussions; Paul Hemeren for useful discussions and comments and, in particular, for his thorough and insightful proof-reading, which at times involved converting my private language into English; Simon Winter for inspiring discussions and for his inestimable and competent help and support regarding all kinds of practical and esthetical issues conceming the appearance and production of the thesis; and Tom Andersson for joining the group and for the inspiration and support he has already given me.

A special thanks to Magnus Haake for the illustrations. When I first saw his illustrations, I knew for the first time what a Be1ry-Creature really looks like.

Also I want to thank all my friends that I have *not* mentioned, for making this thesis possible.

Finally, and most of all, I would like to express my deep appreciation for my advisor Professor Peter Gardenfors. Mere words fall short of capturing my gratitude to him. His constant support and invaluable inspiration throughout the years have significantly contributed to my work. The research possibilities and the constructive atmosphere at the Cognitive Science Department in Lund are to a great extent due to his influence. Thank you for everything, Peter!

One difficulty with a project that extends over several years is that one's attitude concerning research methods may change with time. One may reach a point where one considers starting anew, with the same topic but with partly different methods and issues, *or* continuing with the issues and methods chosen at an earlier point in time. When I arrived at this situation I chose to do the latter. Therefore, in spite of my growing preferences to pursue my own empirical research and also for treating more specific issues within a more limited scope, the present study does not involve any empirical work of my own, and it covers relatively much at the expense of certain details. In this sense, it mirrors my earlier preferences as to ways of conducting research.

PART ONE: THE PROJECT AND ITS FRAMEWORKS

"'Ine first tfiing I've got to ao, " saicf 5t{ice to fierse[f, as slie wanckrea aGout in tfie woocl, "is to grow to my rigfit size againj ana tfie secona tfiing is to fina my way into tfiat {ove{y garaen. I tfiinl(tfiat wif{ Ge tfie Gest p{an. "

It sounaed an e effent pfan, no aouGt, and very neatfy ana simp{y arrangedi tfie oniy clifficu{ty was, tfiat slie fiaa not tfie smaf{est icka fiow to set aGout it...

Lewis Carro{[, 5tfice in 'Wonaer{and

1. GENERAL INTRODUCTION

1. Background

The general topic of this thesis is the planning of action. As a first example of the activity that I have in mind, consider the following scenario. Two friends, Anita and Irene, are on the last day of their vacation together. Tomorrow they will go back home andreturn to work. They are going to leave early in the morning, on a ferryboat. They have finished lunch and are sitting in the shade, resting. Both are feeling tired since they were out very late the night before and have been <loing a lot of windsurfing and swimming during the day. They agree that, as they are both tired, it may be a good idea to take it easy in the evening, have something to eat at the guesthouse where they are staying, go for a nice walk and getto bed relatively early, as they have toget up around six o'clock the next morning. Besides, they have very little money left, and going out in this village is indeed expensive. "Instead", Anita suggests, "we can have a *really* good <linner on the boat tomorrow and even renta cabin." But a couple of hours later, the following discussion takes place:

Irene: Hello my friend, how are you <loing? Oh, just look at that, see that sky, and those little boats! I really love this place. You know what, Anita, I feel like doing something *special* this evening. Are you still tired?

Anita: No, I don't feel tired at the moment, actually.

Irene: You know, I just spoke to our host, and do you know what he recommended? Go to Volerini, an old village 20 miles north from here; there are some splendid, traditional sea-food restaurants anda nice bar where they play live music. Toere is a bus leaving from here in half an hour. It goes along the coast all way, so that you can see the sea ·an the time. Are you hungry? I am, mm, just think of it, I'11 have lobster, crab...

Anita: But, Irene, then it'll be another late night. Shouldn't we think a bit about tomorrow? We have toget up at six, and it isa long journey. How shall we get back from there? Toere is no bus back, I suppose, and so we will have to take a taxi.

Irene: I think it'll be all right.

Anita: But what about tomorrow? Look, *I* think we are going to be very tired. The boat trip will probably be tiresome even if we *don't* go out, and I think it would be a good idea to have enough money so that we could get a cabin, if we want to rest, and I definitely think we should take it easy tonight. I don't think it is a good idea to go out at all.

Irene: I thought you said you were not tired.

Anita: OK, I am not tired *now,* but I know I *wi//* be. We really didn't sleep much last night, and I want to sleep tonight so that I can enjoy the trip tomorrow. I look forward to that, and eating seafood on the boat, like we spoke of. We won't be able to do that if we go out tonight.

Irene: Oh, come on!.. It's such a beautiful evening, I want to see the coast and eat lobster. Come on, enjoy *this* evening *now!*

Anita: I am enjoying it. It's just that, being short of money in this way isa bit worrying. I don't think that it will feel good to leave tomorrow without having much money left. We might really need some. What if something happens?

Irene: You are hopeless.

Anita: But Irene, listen, do you have to rush? Let's sit down and talk about this.
Irene: No Anita, you stay and think! I'm going, and I'm going now. I want to see the coast as long as the sun is up, and I am hungry and want to go to that restaurant. I'm going to get my bag, then I'll come by on my way to the bus stop to see if you have changed your mind.
Anita (shouts after her): Take something warm along, it will get chilly later on.

How does the evening pass? Irene goes to Volerini, and Anita stays. First Anita sits for a while and wonders whether she has done the right thing. She tries to imagine what it would be like if she had gone. For a short while, she thinks about checking to see if there is another bus. Maybe she can go, but only go to the restaurant and not stay afterwards. It would be neither too late nor too expensive. But no, it's hardly worth it. It will take too much effort anyway. It *would* have been different, perhaps, if it were not the case that she has to go to work the day after they get back and that she has an important and difficult meeting to attend. At this meeting, she will present her proposal for a new product. That is not going to be easy. Her boss will certainly be prepared to object to most of what she says. Anita imagines some scenes and wordings. "I will be nervous," she thinks, "but I just have to relax and be sure of what I say; I will go through my ideas again and freshen them up. *When* can I do that then? Tomorrow night when I'm back? I'll probably be tired. In the morning just before will be too late. No, I'll try to do it on the ferry tomorrow." Anita takes a walk on the beach, continuing to think about the meeting, of what can happen and of what she will say. After a while she realizes that she is cold, and that it is very windy. Yes, she really wants to get back now. She has walked quite far without noticing how cold she is. "If it is windy like this tomorrow, I'm likely to get seasick; I'll take some pills for seasickness now when I get back."

Anita returns to the guesthouse. She prepares her things and goes to bed but finds it hard to go to sleep. Her thoughts just keep fluttering around. She thinks about the meeting, the journey tomorrow, and wonders when Irene will be back. Finally, she manages to get to sleep but is shortly thereafter woken up by Irene.

Anita: What time is it?
Irene: Quarter past two.
Anita: Oh no, don't tell me that. Have you had a good time?
Irene: Lovely, just lovely, but I can't tell you about it now, I'm exhausted.

The next day, on the sea journey, both Anita and Irene are indeed very tired. But, it turns out that there are some comfortable chairs on one deck, where they rest and even go to sleep for a while. Then they have a simple lunch on the sundeck and then get some more rest. Anita prepares her presentation, and they talk about their experiences during the week. Irene talks enthusiastically about the night before - and Anita tells herself that if she comes to this island again, she will go to Volerini. In the evening Anita offers Irene a dinner in the restaurant (Irene has hardly any money left). They sit for a long time and talk and have a very enjoyable time at the end of their vacation, talking, amongst other things, about different possible projects and journeys for next summer.

This example shows how two friends are thinking about and attempting to decide *what to do* - and *how* and *when* to do certain things - during the evening, the following day and in the more distant future. One difference between them is that Irene is primarily concerned about her present desires and needs, whereas Anita is more concerned about what may take place later on (even to the extent that she does not notice what happens right now, for instance, that she is getting cold).

One may be inclined to say that Irene's way of planning (and her lack of planning) renders her an advantage over Anita.Butone should note that things could have tumed out differently- more to Anita's advantage and less to Irene's. For example, Irene might not have had such an enjoyable evening, Anita might have gone to sleep and slept long and well, and it might not have been possible to sleep on the boat (or, if they had spent money on a cabin, they would not have been able to have the nice <linner).

The following three observations, which can be related to the above scenario, constitute the background for the thesis:

(1) Toere *isa* certain mental activity that human beings are capable of and frequently engage in; viz: thinking about *what* to do, and about *how* and *when* to do something before <loing it, or, in short, *making plans for one's actions.* A human being can plan a wide variety of actions, for instance, plan what to do the coming evening, plan to run some errands at different places in town in the afternoon, plan next week's work, plan when to wtite a particular letter, plan tomorrow's lunch, plan what to say at a meeting, plana career, plan how toget home from one's hotel from the market place in the city one is visiting, plan a picnic, plan when to do one's wash, make a plan for moving to another country next year, plan to play ajoke on a person, plan the building of a house, plan the use of time during a vacation, plan how to get back the jacket that one has forgotten at a certain place, and so on.

(2) Making plans seems to be a*prevalent* and often important activity in the lives of human beings.[1] At least some individuals appear to be quite *concerned about* this planning of action that they engage in, making utterances such as: "I'm no good at planning but I would like to learn to do it hetter", "This constant bothering about tomorrow and about the future is so consuming", "I really *enjoy* making plans for the future", "Ijust *cannot* plan my time properly", "I always make plans but never carry them out", "I've got a *wounderjillplan",* "When you do something you planned, you feel good", "It is boring when it is all planned in advance and you know everything that will happen", "It's weird - she had planned to seduce him", and so on. Note that these are *evaluative* utterances and notjust 'neutral descriptions'.

The following tension is sttiking. On the one hand it is evident that the capacity to figure out and structure one's future actions is of considerable value. It underlies a great deal of the progress of humankind as well as of achievements of particular individuals, and is an appreciated capability. On the other hand, it seems that the capacity of planning also has negative and problematic aspects. (See the example with Anita and Irene.)

---

[1] The prevalence of planning (in humans or other creatures) depends, of course, upon one's *conception* of planning. Sometimes the planning of action is conceived of as a set of sophisticated cognitive skills, which have to be acquired through instructions and extensive practice and are only applied when one is confronted with certain difficult problems. In that case, planning is *nota* mundane activity that humans engage in to a great extent. On the other hand, planning can be very broadly defined, like, for instance, when planned behaviour is equated with goal-directed behaviour, in which case, of course, planning isa common activity that permeates human action and life. My conception of planning, as we shall see, lies between these 'extremes'.

(3) Some aspects of planning seem to be *uniquely human.* It is, for instance, difficult to conceive of a dog or a chimpanzee that would make a plan for what to do in the evening. The reason why I consider this last point is that I have a separate interest for the issue of *'human nature '*. What - if anything - is *specifically* human, i.e., absent in other species and present in all humans, or, in other words, *uniquely* and *universa/ly* human? Can an understanding of *human planning* shed some light on this issue?

Thus, the background for the thesis is set by the following questions:

• How is the planning of action possible?

• *When* do human beings plan, *how* do they do it, and *why* do they do it?

• Is it possible to trace any evolutionary origins of human planning?

• What in planning - if anything - is *specifically human?*

• What is the *value* of planning? More specifically, what are the positive and negative aspects of planning in general and in human beings in particular?

2. *Purpose of the Study*

The primary aim of the thesis is closely related to the original motivation for it. It is an investigation of the phenomenon of the planning of action in order to suggest answers to some of my original questions and thus to improve our understanding of the role of planning in the acitvities of living creatures, especially humans. I want to do this by situating the capacity for planning in the context of other cognitive and behaviour-regulating capacities. In particular, I want to shed some light on the positive and negative aspects of planning; and not approach it as a 'neutral', purely 'rational' activity. I focus on a single individual's planning his or her own actions and not on an individual's making plans for other individuals or on collaborative planning.

The investigation is not primarily intended for those who are interested in a system whose actions are all driven exclusively or primarily by *plans,* or in a pure planning system, that is, a system exclusively capable of making plans. Neither is the study of immediate interest to those who want a detailed specification of a planning system. The present study is carried out on a more general level. It attempts to present general principles for planning, and it does so at the price of lacking in details. Still it may have indirect relevance for someone who tries to describe or build some kind of autonomous agent capable of planning by pointing to some issues that, in my view, should be considered in this context, for instance: What other activities are supported by elements of a planning apparatus? What are the costs in terms of attention, memory, etc., of the activity of planning? Are there other activities (intemal and behavioural) at the expense of which planning occm-s?

But, primarily, this thesis should be of interest forthose who seek a more general understanding of the role of planning in a biological agent and of the background of the evolution of the phenomenon of planning. Finally, even though the study is carried out on a general level, it may, I believe, give some understanding of certain aspects of *human* planning, for instance, give an account of some kinds of misguided or even pathological planning in humans.

3. Structure of the Thesis

In general terms, this is an investigation of the phenomenon of planning of action from the perspectives of *cognitive science* and *biological functionalism* (which will be presented shortly), where I try to situate planning and planning capacity in a context of some other cognitive phenomena and forms of behaviour control.

A *tentative answer* to the question conceming the uniqueness of the human capacity for planning is that human beings, but not pre-humans, make plans for actions, not only in order to satisfy *current needs* and thus solve immediate problems, but also in relation to *anticipated, potential* needs and corresponding problem situations. For example, the hungry chimpanzee may make a plan for getting the bananas that are not immediately reachable or plan how to get out from a trap where it is caught. But only a human being may, while not the least hungry, thirsty or cold, make a plan where she considers potential, future needs for food, water and shelter, and so on, or make a plan for how to deal with the situation *if* there should happen to be a burglar in the house when one gets home, etc. (like Anita in the example above who is concerned about possible tiredness during the following day's joumey, about having something to eat during the journey, about potential difficulties on a future meeting, etc.). There are pre-humans who are *immediate planners,* whereas only human beings are *anticipatory planners.*

This hypothesis Jormed the basis for some more specific issues that I address in my project, namely, to specify and investigate the distinction between these two levels of planning competence and to discuss the *transition* from immediate to anticipatory planning. The last issue is the problem of how one can, departing from a design of an immediate planner, change it into a system also capable of and motivated for anticipatory planning.[2]

I illustrate *immediate* and *anticipat0ly* planning respectively by describing two kinds of fictitious creatures; the I-Creatures and the A-Creatures.[3] I describe the conditions under which they live, their motivational, perceptual and behavioural functions as well as the nature of their planning. Because of my use of this fictitious illustration, this study, in some respects, has the character of a thought experiment. However, there are empirical considerations underlying the thought experiment. In 'constructing' these creatures, I have been inspired by considerations concerning living biological creatures and by data from ethology, comparative psychology and anthropology. Primarily, the data concerns higher primates, mainly chimpanzees. I also use some anthropolo cal data and hypotheses concerning early humans. The main reason for using a fictitious illustrative case is that I want to be able to discuss properties and occurrences of planning and planning capacities that are, in my mind, possible and plausible, without having to assert that all that

---

2Note that this investigation has an independent interest, even if the hypothesis about anticipatory planning as a hallmark of human planners is mistaken.

31-Creatures are immediate and A-Creatures are anticipatory planners. 'Berry-Creatures' is a general term for both I- and A-Creatures.

I say *is true* of living creatures or that it *has to be true* of *every possible planner.* Furthermore, this strategy is recommended by the need to simplify. Nature is too complex. With a fictive case it is possible to obtain a description that is more manageable. The actual example of the Berry-Creatures could possibly even be implemented on a computer. Finally, the use of an invented case helps to stress the hypothetical character, partly due to the research domain as such, of many of my proposals.

The thesis consists of three parts: Part I, *The Project and Its Frameworks;* Part II, *A Systems Understanding of the Planning of Action;* and Part III, *An Evolutionary Understanding of the Planning of Action.*

In Part I, I present the project and my theoretical and methodological frameworks. These frameworks do not consist of any particular *theories of planning* (although I am inspired by and make use of various such theories) but of some more *general approaches* for dealing with mental[4] - psychological - and behavioural phenomena. The first of my two frameworks is *cognitive science.* As I conceive of it, cognitive science is a *research program* with the central notion being mental representation. The aim of the program is to understand mental phenomena and how these influence behaviour by approaching mental phenomena as instances of information processing (in a wide sense). It is a program that stresses the *constructive* role of cognitive subjects in perception and cognition. My second framework is what I term *biological functionalism.* This approach centers around the theory of evolution and has as its central concepts those of *adaptation* and *goal-directedness.* To investigate a mental phenomenon from this perspective means to aim not only at a *systems understanding* of the phenomenon; that is, to find out *when* it occurs and *why it* (and not something else) occurs at that time, and *how* it occurs in the systems in question. But it also means to seek an *evolutionary understanding.* This implies investigating the following two questions: What are the evolutionary roots and precursors of the phenomenon? And wherein lies its function and selective advantage? I'll have more to say about cognitive science and biological functionalism in chapters 2 and 3.

In Part II - *A Systems Understanding of the Planning of Action* - I start out in chapter 5 with a discussion characterizing the phenomenon of planning and its relation to certain other phenomena such as problem salving, dreaming and intending, etc. On the basis of this, I come up with a *general definition* of planning. I also define and discuss the two levels of planning that I have mentioned; immediate and anticipatory planning. In the first chapter I also introduce the Berry Creatures and their World. In chapter 6 I sketch a design of a basic *perceiver-and-behaver not* capable of planning. I present my views on perception, introduce the idea of behaviour as the control of perception and discuss perceptual-behavioural systems of different orders. In that chapter I also present many elements that will be useful when endowing such a basic perceiver-and-behaver with a planning capacity. Chapter 7 is called 'How to Obtain Planning'. In this chapter I give some examples of planning. I explain why the perceptual-behavioural system cannot

---

[4] By 'mental' - which I use interchangeably with 'psychological' - I mean the following. Mental states and processes are those states and processes appealed to in psychological explanations, both those characteristic of current behavioural and cognitive psychology and 'mentalistic' explanations in philosophy of mind. In particular, 'mental' is not used as an opposite of 'physical' and it is undiscriminative between conscious and un- or sub-conscious. It is also *not a priori* related to 'biological stuff': nervous systems and the like.

plan, I introduce the capability of *imagination* and then I introduce planning as a form of *imaginative generating-and-testing of mental representations.* Furthermore, I discuss how and why planning occurs at the expense of behaviour and perception and finally there is a discussion of generate-and-test models of planning. In the next chapter, called 'An Immediate Planner', I try to address the issues of 'when', 'why' and 'how' planning occurs in an immediate planner; that is, to give a systems explanation of immediate planning. In chapter 9, 'An Anticipatory Planner', finally, I first discuss the distinction and transition between the two levels of planning capabilities and then attempt to give a system explanation of anticipatory planning.

In part Ill, which is the final part, my aim is to contribute to an evolutionary explanation of immediate and anticipatory planning respectively. Chapter 10 deals with the evolutionary roots of planning. I discuss three stages of behaviour control; the instinct system, the trial-and-error system and the playing system. Chapter 11 concerns the value of planning. Chapter 12 is the most speculative. Here I attempt to explicitly discuss *human* planning.

There are two features that distinguish this thesis from many other studies of planning:

(1) The emphasis in the investigation lies as much on the question *'why* planning?' as on *'what* is planning?' or on *'how* does planning work?' (*'how* does one plan?'). This stands in contrast to a more philosophical study which would focus more on *what planning is;* i.e., on what is *constitutive* of planning and on how the concept of planning relates to certain other concepts. It also stands in contrast to an AI-perspective where the primary aim is to *construct* a program that *does plan* - and often nothing but plan - thus, focusing on *how* planning works. From neither of these two perspectives does the question of 'why planning?' assume a central place.

'Why planning' divides into two topics. *First,* for a given kind of system, I ask *why* the system plans *when it does:* Why, at certain particular moments, does the system *start* to plan and *continue* to plan?[5] This question asks for the driving forces, or motivations, behind planning as an activity. The original reason why I started to think about this was that I often met people who complained that they were planning too much. This led me to ask about what it is that drives or motivates planning in an individual system. And indeed, when one takes seriously the fact that all systems have only limited resources and combines it with the assumption that not only extemal but also *internal activities* compete - both with one another and with extemal activities - for those resources, it is clear that this question of motfvation must be posed for *internal activities* as well as for *behaviour.*[6] *Second,* I consider the question 'why planning?' from an evolutionary perspective: *Why* are there planners at all, and why planners with a specific planning capacity? What selec-

[5] In other words, what sets off, sustains and reinforces planning as an activity?

[6] Like asking what motivates an individual at a particular moment to attack instead of escaping or looking for food, or what motivates an individual at a given moment to start searching for food instead of going to sleep.

rive advantages does a planning capacity - or a particular planning capacity - render planners? *For what good* is planning, or, in other words, what is its *evolutionary value?*[7]

(2) A second aspect in which this study differs from many studies of planning is in its use of a blend of hypothetical or invented and empirical material. On the one hand, the descriptions of planners and planning that I give and the explanations that I suggest are *not* based only on empirical material and observations. I have not observed intemal representations of goals, mental generating-and-testing, anticipatory processes, and so on; I have not observed the emergence of anticipatory planners in nature. On the other hand, it is not the case that I atttempt to produce just a description of *any* planner, as long as it is coherent and works. I *do* engage in construction in a sen se, namely, in the 'construction' of the 1-Creatures and A-Creatures. Yet, the *central* hypotheses, descriptions and explanations I put forward are only those that I find *plausible* in the con text of current research and knowledge of the biological creatures that are planners. In other words, the descriptions are not to be taken as abstract descriptions of all possible planners, but they include several assumptions about what I take to be planners in nature. My *prime* aim is to hetter understand planning and the development of a planning capacity as a *biological phenomenon.* This does not exclude, however, that some of the proposals that I make may be useful for someone who wants to construct an artificial planner.

[7]There is more than one perspective from which to consider the *value* of planning. Whereas valuable from an *evolutionary perspective* means that there is a *kind of design* that in a sufficient number of cases renders individuals with this design a selective advantage, one may also speak about the value of a planning capacity, or of a particular instance of planning, from a *particular individual's perspective* - *reflected* as well as *experiential* value. Take, for instance, the question of why an individual is making plans for a particular voyage, or a holiday. Reflected value is illustrated by answers such as 'I gain time thereby' or 'I reach my goals hetter'; and experiental value by answers such as 'I enjoy thinking about and structuring my holiday' or 'I get satisfaction from having things under control and knowing what will happen'. One may also speak about the value from a *social perspective* and find assertions such as the following: 'Planning is valuable because it makes it possible to coordinate actions and collaborate and thereby reach collective goals', 'The planning of action can be valuable because it may make one think about what other individuals want', 'Planning *together* can be valuable because it leads to group cohesion, makes actions run smoother and involves valuable social experiences (getting to know each other, and so on)'.

2. COGNITIVE SCIENCE AS A FRAMEWORK

1. Introduction

The main framework for this thesis is that of cognitive science. This is an interdisciplinary resarch program that has taken shape primarily during the last fifteen years or so. It is the study of *mental processes* in knowledge handling: processes in perception, imagination, reasoning, remembering, recognition, learning, planning, language production, language understanding, and so on. Cognitive scientists construct and investigate models of mental activities and try to understand how various mental activities relate to one another as well as how they relate to *behaviour.*

At the center of cognitive science one finds parts of cognitive psychology and Artificial Intelligence (Al) as well as certain topics within philosophy, linguistics, anthropology and the neurosciences. Cognitive science can be said to have emerged as it 'became clear that various research traditions were dealing with certain common issues concerning cognitive processes.

2. Two Key Features

I see *two keyfeatures* in cognitive science: (1) the idea of mental representation and (2) the idea of psychological functions.

2.1. The centrality ofthe concept ofmental representation:

Mental processes are approached as instances of *information processing,* or as transformations of *mental* or *internal representations.* Such internal representations are, in turn, assumed to be *instantiated* in the brain, the nervous system and in neural processes. It is difficult to overestimate the importance of the notion of mental representation. Toere are indeed cognitive scientists who devote a lot of energy to the concept as such and attempt to develop more or less general theories of mental representation. I will not present such a theory, but only give a brief characterization of the notion as I conceive of it.

In my view, the best way toget a grasp of the notion of an *internal representation* is to conceive of it as an internal parallel to an *external representation;* that is, to think of an internal representation as some kind of intemal map, scheme, symbol, etc. Let us consider a map of a certain forestial are_a, used for cross-country running. Using this as an example I will list some properties that are characteristic of representations, external or prototypical representations (maps, drawings, scores, photos) as well as internal. But I will start by introducing two important distinctions. First, there is the distinction between *representation and content.* The map represents a particular area in the forest. It is possible to represent this particular area in other ways, for example, to take a photo of the area, to use other kind_s of maps with other notations and so on. Such representations may be transformed into each other, more or less successfully. The *same content* gets *re-presented* in - more or less -*different ways.* Second, there is the distinction between *representation* and *medium.* The map with its symbols may be printed with different kinds of printing ink on different kinds of paper, it may be drawn with pencil, and it may in principle be realized using some cloth and some threads. The *representation* is the *same* but the *medium differs.* (See Glass and Holyoak, 1986, pp.5-9.)

A representation is a format for presenting a certain kind of information. The following are properties of representations.[8]

(1) *A capacity for correlation:* A representation has the property of 'standing for', 'pointing at' or 'being about' something, namely *the represented,* and it may represent this more or less correctly. Toere can be a correlation between (some of) the properties of the representation and (some of the) properties of the represented objects. (We find two small streams represented on the map, just as we find two streams in the forest in question.) This correlation of course is no accident. It is often the case that if the represented object had had certain other properties, then the representation would also have been different. (If there had been three small streams in the forest, three small streams would have been represented on the map.)

(2) *A capacity to mis-represent:* A representation has a capacity of being erroneous. It can represent non-existing objects or attribute non-existing properties to existing objects. (A hill may be represented on a map although there is no hill in the actual area or a hill that is there may be represented as higher than it is.)

(3) *Compositionality:* Representations can have parts that may themselves be representations. (On the map there are lines that compose curves, a group of curves form the representation of a hill, and the whole map as one representation is composed of parts that represent different areas and entities.) Representations often have a hierarchical structure, and hence there can be different levels of detailedness in a representation.

(4) *Scope:* A representation has a certain domain of application or a certain reach. It makes explicit *certain kinds* of information. (The kind of map that we are considering as an example represents information about spatial relations, distances, heights, hills, rivers, and so on, but not information about buildings and roads and monuments or about sounds and tones, and also not information about the animals that live in the forestrial area or about the acidity of the water in the area.)

(5) *Selectivity:* Scope implies selectivity, but there is selectivity also within a scope. A *particular map* of the kind we are considering represents information about *one particular area.* It has what is sometimes termed *a proper object.* (Scope is a property that applies to a *kind* of representation, like a kind of map, whereas selectivity also applies to a *particular* representation, like a particular map.)

(6) *Perspective and focus:* No representation will represent all properties of its proper object, but it will select certain properties. It will take a certain perspective. First there are the *spatial and temporal perspectives.* The object is represented from a certain position and at a certain point in time. The forestial area is represented from above rather than from below and at some point in time that is recent rather than hundred years ago. Second there is the *perspective of purpose or interest.* Two representations with the same scope and the same proper object (same selectivity) may *focus* on different aspects. A map that shall be used for cross-country running at night may differ from one that shall be used for

[8] Tue list is based on Dan Lloyd (1978). It is not to be taken as a *definition* of a representation, as not *all* of these properties are *necessary* for something to be a representation.

cross-country running during the daytime. And these will differ from a map that shall be used for purposes of charting mineral deposits in the area. Focus also has to do with which information is made more explicit and which is more implicit. Distances between the different controls in a cross-country run and start may be explicitly written on the map. But if they are not, one can get at this information by measuring and using the scale on the map. How explicitly some information is represented can only be decided strictly relative to the representation *and* the processes and functions that are available for using it.

(7) *A capacity for being used:* If the first characteristic (a capacity for correlation) implies that a representation is a *representation of something,* this last characteristic implies that a representation is a *representation for something/someone.* It is a representation for a *user.* For instance, a map may be filmed and thus represented on a film. To get at and use this information, the film must first be replayed on a projector (a cassette recorder, for instance, will not do). Furthermore, the system that is presented with the film must be one that can see and that can read maps of this kind, in order to interpret and use this representation. Or, to consider another example, a representation in Braille is a *representation for* a user that knows Braille.

I have here used the example of a map, but this does not mean that I am saying that *mental representations* must be map-like. One can conceive of internal representations that are like pictures, like words, like schemas or like mathematical representations of some kind. Also, nothing is required about the awareness of mental representations. I use the notion for covering a whole spectrum: *from* implicit, hardwired representations that do not and cannot involve processes of which an organism is conscious or aware (for instance, the representations of physical information that, according to Marr (1982) are created during the early stages of visual processing) *to* more flexibly programmed representations involved in problem solving and classificatory activities that organisms carry out with some degree of awareness that they have created and are using a mental model. (See Gardner, 1987, p.384.)

The following idea is crucial: the level of mental representation, as a level of analysis, is *in priciple independent* both of the *neurological* level on the one hand, and of the *social and cultural* level on the other hand. It is independent in the sense that there are questions about human mental activity and behaviour that can be posed and approached on this cognitive level without dealing with and knowing anything about the neurological or the socio-cultural levels. This does not exclude that research from ny of the three levels can shed light on and constrain research on other levels, and that ideally a study of human psychology and behaviour should deal with all three levels. But the point is that it is *useful to posit this cognitive level of analysis* in order to gain an understanding of human mental activities and action. It is too limited to try to explain human behaviour only with reference to facts and entities on a neurological-biological level on the one hand and a social-cultural level on the other. For instance, in order to answer, say, why two different individuals behave differently in a situation it may be insufficient to analyse this in terms of biological-neurological differences and/or differences on a social-cultural level (like different education, different language and culture, different personal history, and so on). Sometimes the most appropriate explanation refers to differences in their mental representation of the situation.

What role does this first key feature of cognitive science play for the present study? First, in several definitions and discussions I rely upon the concept of mental representation. Not only do I conceive of *planning* as an operating on internal representations, but I believe that intemal representations play an essential role in other kinds of behaviour control as well. They enter already into my sketch of a basic behaving-and-perceiving system that is not capable of 'higher cognition' such as planning. In my view, *perceptual representations* are the most basic mental representations and the source of non-perceptual mental representations. Second, the idea that two systems in the same external environment or situation may *represent* the situation differently (perceptually and cognitively) is important. The plan a system comes up with - and how it realizes the plan - depends on its representation of the situation. Let me give some examples:

• Two individuals are both hungry, and both are going to be tired relatively soon. Butone represents the situation only as a situation where there is no food around and the other represents it as a situation where there is no food and no suitable place for sleeping, although it will soon be tired. These individuals may produce different plans for action.

• Two individuals plan to get some tools at a certain site and thereafter go to the site where the tools are going to be used. One individual represents this problem in terms of how far he will have to walk, another focuses on how far he will have to walk with the heavy tools.

• A bridge is broken. One individual represents this as a board that is missing, another individual represents it as a piece of wood that is missing, and a third represents it as just *something* missing that means that you cannot pass over the bridge. These individuals may also think differently when trying to figure out how to deal with the situation.

Third, let us return to the idea of the independency of the cognitive level as a level of analysis distinct from a socio-cultural level as well as from a neuro-scientific level. In the main part of this thesis I do deal with issues conceming planning without bringing in any social or cultural considerations, for instance, issues like the following: 'Departing from an immediate planner, what must be changed in order to make it into an anticipatory planner?', 'Do immediate and anticipatory planning involve representations of time, and if so, how?' 'What kinds of change does a planner have to know of about and represent?' And the study is pursued with only marginal references to neuroscience.

However, at some points I *do* bring in socio-cultural aspects, and even though my use of neuro-scientific studies and data is limited, I (try to) use 'higher-level biological constraints', taken primarily from ethology and from evolutionary theory. And it is clear to me that for a full investigation of the role of planning in human activity, these levels and aspects should be fully involved. Such a study would be a study that integrated approaches from psychology, anthropology, neuro-science, (and AI and philosophy), employing the notion of mental representation as a central notion.

2.2. The centrality of the idea ofpsychologicalfunctions

The second key feature of cognitive science, in my view, is the idea of *psychological functions*. Important roots as to what is sometimes called *psychologicalfunctionalism* can be found in the ideas of the American psychologist and philosopher William Jarnes (1890). James stressed that different psychological mechanisms exist because they help individuals to carry out important activities. Our various ways of *thinking andfeeling* are *useful* in shaping our reactions to the outer world.

Psychological functionalism encourages a study of *mentalfunctions* - in perceiving, recalling, thinking, dreaming, imagining, planning, and so on - that are *relatively stable* and recurring even though the specific contents vary. (This is in contrast to psychological structuralism that focuses on the relations between specific contents - content pattems or structures.) It isa study of functions, sub-functions (the composition of functions) and of how they relate to each other. It also involves a certain focus on mental functions that are common to all human beings rather than a focus on individual and cultural variation and on specific mental phenomena in specific individuals.

These ideas are of relevance for my work. I approach planning as a general competence, applied to various contents and as a function that shapes behaviour, but I say very little about individual and cultural differences concerning the capacity of planning.

The growth of these both key notions of cognitive science - that of mental representations and that of mental functions - has, I believe, been helped by the appearance of computers. Even if cognitive science is *conceivable* without the computer, the development of computers *actually* has been an important factor in the rise and development of the fl.eld. As Ulrich Neisser puts it:

> [...]the activities of the computer itself seemed in some ways akin to cognitive processes. Computers accept information, manipulate symbols, store items in "memory" and retrieve them again, classify inputs, recognize patterns, and so on. Whether they do these things just like people was less important than that they do them at all. The coming of the computer provided a much-needed reassurance that cognitive processes were real; that they could be studied and perhaps understood. (Neisser 1976, pp.5-6.)

Because the computer is a physical system - like the brain and nervous system - where such processes seems to occur, intemal processes of this kind no longer seemed so mystical and inaccessible to investigation. The use of computer models as a *generalframework* for theorizing about mental phenomena and behaviour is common in cognitive science, as well as the use of computers as a *tool* for modelling mental processes. A common strategy, in particular in AI and AI-related research, is the following. You try to formulate a theory of the phenomenon that you want to understand, which is so detailed that it is possible to write a program that will *instantiate this theory.* Running such a (computer) program, that at some level seems to work according to similar principles as those guiding the actitivy (function, capacity) being modelled, is said to be a *simulation* of (aspects of the) activity. In a simulation one may note similarities in inputs and outputs, order of steps, relative speed, reactions to various disturbances, etc. An example of this kind of research is Hayes-Roth and Hayes-Roths' model of planning (1979). They first studied human subjects given a planning task of scheduling a series of errands in a hypothetical town. The subjects were required to think out loud while planning, and a transcript, or protocol of this was taken as evidence of planning operations. They then constructed a computer program which reproduced human performance (by maintaining similarity in the depth of search, requiring a similar amount of time, by making similar mistakes, etc.)

Although the fundamental concern in cognitive science is with *human* mental processes and *psychological reality* ('What happens in humans and how?', 'How is the psychological phenomenon p and behaviour b possible in human beings?', 'What are the underlying structures and processes?' etc.), there is a parallel dealing with questions like: 'How is the psychological phenomenon p or the behaviour b - or something interestingly similar to it - possible *at all?* 'The hope is to construct a program that does p or b, or seems to do p or b or, at least, does *something* that sheds some light on how human beings perform p or b.[9]

3. Cognitive Constructivism

An idea that is central to cognitive science and closely related to the notion of mental representation, is that of *cognitive constructivism,* with its key concepts *selectivity and interpretation or subjectivity.* (Cf., for instance, Boden, 1977; Clark, 1986; Oatley, 1978.) To present this, let me start from the notion of *information.* Information, as I understand it, is basically an issue of a *choice between alternatives.* There *is* information whenever a given system is in one possible state rather than in another, and information has been *transmitted* when the state of one system is in some way dependent on the state of another system. At any point and in any situation there is more information than any cognitive system[10] can store, encode and deal with, as every cognitive system is finite and has limited resources.

Therefore, an information-processing system or a cognitive system has two fundamental or basic problems to deal with: (1) to *select* among the information available and (2) to do this *in such a way* that the selected information is *coherent* and *useful* for the system; i.e., it should be related to the problems it has to solve. *Constructivism* says that this selection is directed by *pre-existing informational structures* in the cognitive system whether these be innate or learned. Mental activity - perception and cognition - is basically a question of *construction.* Mental processes are *active processes* directed to construct sense - to *'make* sense' - out of a constant flow of information in the complicated surrounding world. On the basis of changing masses of information which are at times incomplete and degenerated *the system constructs* representations or models of the world around it. And there is an interplay between information in the world and information in the perceiving system.

There is construction *all the way,* from perception to higher cognitive functions. *Perception* is held to be the most fundamental *cognitive and constructive* activity upon which all other mental activities build.[11] *To perceive* does not mean simply 'to read off'

---

[9] Sometimes a distinction is made between *psychological* AI and AI (or between *cognitive simulation* - CS - and AI). This distinction relates to whether the research *aims at* modelling or simulating cognitive capacities 'found in nature', in particular *human* capacities and performance, or whether it *aims at* the construction of intelligent programs - by whatever means - with no interest in modelling or gaining insight into natural (human) competencies. In psychological AI, the main aim is thus to improve our understanding of human psychology by devising computational models of human competence. It is only psychological AI that can be said to be a central part of cognitive science. (See, for instance, Clark,1986, and Dennett, 1978.)

[10] By a *cognitive system* I mean a system that can (1) receive, (2) encode, (3) store, (4) retreive and (5) use information. All these activities require *representations.*

[11] For an alternative view take that of J. J. Gibson (1979). Gibson holds that *perception is 'direct';* that there are higher order structures in the 'ambient light' that *directly* specifies the objective properties of our environment, and are *not* cues or indications which need any mental supplementation, compensation or

percepts from the world of external stimulation. There is *interpretation* in perception, whereby the (cognitive) system attaches *subjective significance* to information originating from the external environment, treating this information as *cues* to matters of its own *interests*.

In general, *meaning* is not detected but constructed. In the absence of cognitive systems actively imposing their own constructive schemata on the input of the outside world, there would be *no cues, no meaning* - no perception or cognition at all. Note that this relates to the issue of *selectivity*. Two different cognitive systems may use a certain informational input as a cue for *different* aspects of the world, where this is due to differences in the rest of their informational structures, on what their goals are, what their interests are, etc. Because of the richness of the world, the variability of problems to solve, and the many possibilities of doing various things in the world, there is always more than one way to *make sense* of any given input (of a situation, an event, etc.). The stress on the contribution of the subject to cognition and behaviour is, in my opinion, central to cognitive science.

The cognitive constructivist view of perception is important for my conception of the capacity of planning as building upon many aspects of perceptual capabilities. Of importance is also the idea that two systems in the same extemal environment or situation may represent this environment differently (perceptually and cognitively). What plan a system comes up with - and how it realizes the plan - depends on its representation of the situation. (Cf. p.14.)

4. The Philosophical Aspects of Cognitive Science

I find two reasons for saying something about the philosophical side of cognitive science. First because my study, as many other studies in cognitive science, contains certain philosophical questions as well as background assumptions. The second and more direct reason relates to my use of ficititious creatures and thought experiments in this study.

Cognition literally means *knowledge*. And cognitive science may be characterized as an *empirically based study of knowledge;* of its components, sources, organization, development and deployment. In this, the roots of cognitive science go back to issues in epistemology and the philosophy of mind. Several questions that cognitive scientists deal with today seem to have such roots. One asks things like what it means to know something, what it is to be mistaken, where knowledge comes from, how it is stored and used, how it is represented, how different modes of representation relate to each other, what imagination is, how it relates to perception, what an image is, what a concept is, etc. One seeks to understand the constitution of knowers; their perceptual apparatus, mechanisms of learning, memory, rationality, and so on. Other issues that appear are those of the relation between cognition and emotion, of what is innate and what is learned, what is culturally bound in human cognition (what - if anything - is uniquely and universally human). The cognitive science approach to these issues, however, is *interdisciplina,y* and, most importantly, *empirically based.* (Cf. Gardner, 1985, pp.4-6.)

---

construction. He also maintains that perception is totally separated from functions such as anticipating, planning, and remembering.

Not only has cognitive science roots in philosophy, but philosophy is still an integral part of it. Philosophers interested in cognition discuss fundamental questions underlying the enterprise, analyse the questions that are asked, help to put them more clearly, and so on. However, it is not philosophy in general that is this integral part, but roughly what is called naturalistic philosophy; in particular, naturalistic philosophy of mind (with naturalistic epistemology as apart). I take a naturalistic philosophy of mind to accept the following:

• *Mental phenomena* are part of the *natural order.* Mental processes are processes in *material* systems, and sciences such as evolutionary biology and evolutionary and developmental psychology will, in the end, account for the phylogenetical (evolutionary) and ontogenetical development of mental competences and performances.

• There are *constraints* on what *kinds of explanations* a naturalistic approach to psychological phenomena may employ. Matter is the only permitted stuff. There may not be any non-evolutionary teleological explanations. And the explanations may - in the end - not appeal to any homunculi, but all processes must be realised by mechanistic means.

The naturalistic philosopher of mind is concerned with questions regarding the *structure, function and origin of mental phenomena within those naturalistic constraints.* She is indeed concerned with much the *same* questions as the (natural) *scientists* of the mind - developmental psychologists, AI-researchers, psycholinguists, neuro-scientists, etc. Yet there are aspects of this investigation of mind that are what I want to call *philosophical.* I'll try to explain this. Whereas the (natural) sciences aim at finding out *what is* in the world - and how it is - philosophy, as I conceive of it, aims at finding out *what is possible* - and how it is possible. And in dealing with questions of mental phenomena, where there is, at present, so little consensus about answers to, and on how to approach issues of the kind 'what is thinking/planning/imagining, and how does it work?', there is a need for dealing with questions of the kind 'what can thinking/planning/imagining possibly be, and how can it possibly work?' A particular task for the naturalistic philosopher is to try to establish that *naturalistic* solutions are *possible* in principle. And a good way of doing this is to actually *come up with a solution* - even if it is a fairly *sketchy* and *speculative* one.

When actually designing such solutions and attempting to supply some detail to them, the naturalistic philosopher will also use various *empirical* constraints given by what natural science tells us to be the case, *but* as she usually is more aware (than the scientist is) of the status of those various constraints, she is able to 'experiment' with them, combine them and also *question* them - while remaining within the naturalistic framework. This kind of thinking - hypothetical thinking, questioning the essential role of various aspects of a phenomena and taking as little as possible for granted - is what I foremost associate with philosophical activity. My study is to some extent characterized by a balance between the use of empirical data on the one hand and fiction on the other, in particular, in my use of the illustrative case of the Berry-Creatures.

3. BIOLOGICAL FUNCTIONALISM ASA FRAMEWORK

1. Introduction

What I call *'biologicalfunctionalism'* is centered around the (Neo-)Darwinian *theory of evolution* and the concept of *adaptation. Adaptive behaviour* is behaviour that enables organisms hetter to survive and keep fit and propagate. Roughly, *genes* that produce or contribute to the development of physical structures (physiological, neurological, chemical, etc.) that, in turn, produce or contribute to adaptive behaviour will themselves propagate. The driving force behind evolution and the 'ultima ratio' of all organismic existence is the following: to leave as many duplicates as possible of one's genes in the next generation, or, in other words, to *maximize one's inclusive fitness*.[12] (See, for instance, Maynard Smith, 1978; Ruse, 1979; von Schilcher and Tennant, 1984.) A (type of) behaviour that produces or contributes to an increase in inclusive fitness and any structure, mechanism, capacity, etc., that produces this behaviour is selectively advantageous. It has what I call *evolutionary value.*

All species, man included, are products of evolution, that is, of mutation and selection. Through evolution, a species - its physiology, anatomy, motor capacities and its motivational, perceptual and cognitive mechanisms (its ways of thinking, perceiving, etc.) - gets genetically adapted to its environment. The following points about adaptation should be noted:

(1) Ultimately it is only *behaviour* that can be adaptive or not. Structures, functions and capabilities are adaptive (or have evolutionary value) in virtue of the behaviour they produce or contribute to.

(2) The nation of adaptiveness also applies to *psychologica/* competences - like a certain planning competence - insofar as these are grounded in physical structures in the brain and nervous system, and influence or structure behaviour.

(3) Whether and how a particular aspect of a system is adaptive for the system isa function of:

- The biological problems the system has to salve as posed by its constitution;
- The ways for salving these biological problems that the environment offers as well as the obstacles that it poses;
- The rest of the system's possibilities for salving these problems.

(4) In order for a structure, function, or behaviour pattem to *be adaptive* - or to *have evolutiona,y value* - it is not necessary that it is so in all instances. It is sufficient that it is adaptive in a certain critical amount of cases. (This applies both if 'cases' means different individualsofa species and if 'cases' means different occasions in the life span of an individual.)

---

[12] The proliferating of one's own genes, note, does not have to occur (exclusively) through one's own offspring but may also occur through other relatives.

(5) Adaptiveness or evolutionary value is always related to a certain *environment* with its tasks and opportunities. (Cf. (3) above.) Thus, a kind of behaviour, a structure or competence may *have been* adaptive but have become non-adaptive or maladaptive because of changes in the environment. Phylogenetic change is slow. A behaviour pattem in a speeies or individual that is maladaptive or indifferent in a current environment may have been adaptive under the circumstances during which this behaviour pattern was formed and selected. This concems not least the domain of human cognition.

(6) Any capability can be used in other ways and for other tasks than those it was selected for.

Points (4), (5) and (6) above indicate two explanations of how there can be mal- or non-adaptive evolutionary products. First, a celtain capacity, behavioural pattern or mechanism does not have to be adaptive in all instances in order to be selected. It is sufficient with *overall* adaptiveness. Second, there is the possibility of 'more serious' non-adaptiveness, if there is an important environmental change relative to when the capacity, behaviour pattern or mechanism was selected. I will discuss these points both in the chapter on the evolutionary value of planning and in the chapter on human planning.

Applying the approach of *biologicalfunctionalism* to biological - inclusive of psychological - phenomenona means to search both for a *systems explanation* and for an *evolutiona,y e.xplanation* of the phenomena.

A *systems explanation* of a phenomenon such as the phenomenon of the planning of action involves answering the following questions:

• *When* - i.e., under what circumstances - does the system *plan?*

• *Why* does it *plan* in these situations? How does planning fit in with other activities in the system? (This is the question of *'why planning?' from the system's point of view:* 'why plan now instead of doing something else?') In other words: how is the system *motivated* to plan in these situations?

• *How* does it *plan? What* are the mechanisms and processes involved, what kind of knowledge and what kind of competences does the system have and use in planning?

An *evolutiona,y explanation* of a property of a system involves answering the following two questions about the property:

• What are the *evolutionary roots,* the precursors, of the property? What mechanisms does it build upon?

• What is the *evolutionary value* of the property? What adaptive function does it have or has it had? What problems does it or has it solved? (This is the question of *'why planning (systems)?' from the evolutiona,y point ofview.)* [13]

13The question of *'why* a particular phenomenon?' is indeed multifaceted: one may ask for immediate causal factors but also for the developmental (ontogenetic) history of the phenomenon, for its evolutionary history and for the *value* of the phenomenon from various perspectives.

To expand on the issue of evolutionary roots, consider some further features of the evolutionary process:

(1) In evolutionary change there is a great deal of continuity. Evolution proceeds in a *gradual* and *step-wise* manner, aggregating *many small* changes. Complex fine-tuned functions or mechanisms cannot be worked out in advance, but are only likely to evolve as results of an aggregation of small changes made in simpler forms. (In particular, *complex* systems and capacities that evolve are aften *hierarchical.* Already existing stable structures, performing their own tasks - *modules* - are *combined* into new structures.) (Cf. Simon, 1969; Rozen, 1976; Fodor, 1983.)

(2) Evolution 'makes use of' what is already present in existing organisms. It does not come up with wholly new complex structures and mechanisms, but it consists in combining and differentiating existing abilities and extending existing capabilities to new domains. (See Clark,1986, pp.53-54.)

Forthese reasons it is meaningful, when studying a certain phenomenon (a function, a capacity, etc.}, to also investigate its evolutionary precursors. Insofar as the *human psyche* is shaped by evolution, this means that we can gain insight into (aspects of) it by investigating its origins. Human planning of action as an advanced form of behaviour control does not just emerge out of nothing and thus has its own mechanisms and stmctures, but rather builds upon other forms of behaviour control. It can be assumed to have roots in simpler forms of planning that, in turn, have roots in more primary cognitive functions and forms for behaviour control (Cf. Clark,1986, pp.53-54.)

A last note on this topic: When speaking of *genetic adaptation,* one may think in the first place of stmctures, functions and kinds of behaviour that develop in individuals *independently of* any *specific* experiences that the individual has that could give the stmcture, function or behaviour pattern its particular characteristics. But besides this, it could be ar gued, there is also *adaptation* that is *not genetic* but that *does* build upon specific experiences and learning in an individual. Like, for instance, when a child that grows up in China adapts to its situation and certain problems it has to salve by developing a capacity to speak and understand *Chinese.* However, it is not the case that these are two different kinds or categ01ies of adaptation with different underlying mechanisms. It is not possible to divide ontogenetic development - and thus adaptation - into genetically regulated adaptation versus adaptation regulated by learning and particular experiences. A *capacity to leam* and to adjust to specific experiences and problems that have not been foreseen by nature. is not itself necessarily independent of genetic control. There are many forms of interplay. For instance, the development of a particular behaviour pattern can demand specific experiences in the sense that one has to learn this behaviour by imitation and training - but there can be a genetically regulated inclination to be attentive to or even 'search for' the behaviour and to imitate it. Even behaviour that appears to be totally dependent upon specific experiences and the environment can have some form of genetic basis. And, on the other hand, even extremely hardwired ('programmed') behaviour mechanisms and patterns, like refl.exes, may be disturbed by extreme environmental conditions. In sum, the most adequate description is, I believe, that there is a *spectrum* of behaviour from what is *greatly unaffected* to what is *greatly affected* by specific experiences and by particular environmental circumstances. (Cf., for instance, Lorenz, 1973, pp.112-143.)

How does *biological functionalism* relate to *psychological functionalism,* which I spoke of in the previous chapter? It is not a straightforward relation. First, psychological functionlism *only* concerns mental or psychological phenomena, whereas biological functionalism has a broader domain. In principle biological functionalism has some implications for everything that is grounded in some physical substrate and whereupon there are some genetic influences. Furthermore, the psychological functionalist can limit his studies and claims to *human* psychology. There one finds the idea of purposiveness, the ideas that our psychological mechanisms are *useful* in some way, but, in principle, the psychological functionalist need not take an evolutionary perspective and he does not have to deal with sub-human psychology. For biological functionalism, on the other hand, the evolutionary perspective is essential, and there is an inherent interest in comparative studies. Finally, with biological functionalism there is more focus on studies of *behaviour* and of the mental realm as behaviour-regulating, whereas the psycholological functionalist may stay more within the mental realm in his investigations.

But also note that *in principle* one can be a *biological functionalist* but refrain from applying this approach to *mental* phenomena because one does not believe these to be part of the realm that is shaped by evolution. However, as I am dealing with a mental phenomenon - the capacity for the *planning* of action - I am only interested in the tradition that *does extend* the evolutionary perspective to psychological phenomena. The first proponent of the tradition was, as I see it, C. Darwin (1872) himself. Two further sources of inspiration are K. Lorenz, one of the founders of ethology who seeded many ideas concerning comparative studies of behaviour and psychology, and N. Bischof, dean of the Department of Psychology, biomathematical section, at Zürich University, who has cultivated some of those ideas. I also want to mention as a main proponent of this tradition, J. von Uexküll, whose path-breaking influences, even if they were indirect, have been of great importance. Other researchers belonging to this tradition are I. Eibl-Eibesfeld, E. von Holst, D. Premack and N. Tinbergen.

2. *Subjectivity and Spontaneous Behaviour in Biological Organisms*

There are some *conceptions of living organisms* that characterize the tradition of biological functionalism (with proponents as Lorenz, Bischof, von Uexküll). Organisms are in general seen as *subjects* that are capable of at least some *spontaneous behaviour.* Hence they are not conceived of as purely *re-acting objects.*

The first element, the *subjectivity,* originates from *internal structures in* the individual that impose order on the organism's perception of and interactions with its environment. These render the organism a certain autonomy in relation to its environment. In the first place, the individual is endowed with these internal structures because it belongs to a particular *species* and secondly, at least for higher animals, also because of its specific experiences. Organisms that belong to different species may live in one and the same external physical environment. Yet they will perceive it and interact with it very differently. Following von Uexküll (1970), an organism lives and behaves in its particular *Umwelt.* The Umwelt of an organism - as it belongs to a particular species and, to some extent, as it has had its own particular experiences - has both a perceptual and an active aspect. The perceptual aspect - the 'Tastraum', the 'Sehraum', etc. - is what the organism can perceive. The active aspect - the 'Wirkraum' - is what the organism can do with and in its environment. Both aspects - the perceived environment as well as what is done in and with it - differ substantially among different species. In other words, there is species specificity both on the perceptual side and the motor side. A 'parrot thing' is radically different from a 'dog thing' anda 'human thing' even if it is one and the same physical object. Anda living-room is very differently perceived bya human, a dog anda fly. (See the illustrations in von Uexktill, 1970, pp.58-59.)

The second element is that of *spontaneity.* Organisms are not only *re-active;* passively *waiting* for behaviour-releasing input or stimuli upon which they then act in the pruticular ways that are characteristic for the species. But they also sometimes *actively* and without any external triggering event *search* for certain events or situations. W. Craig, in a famous article from 1918, introduces the terms *consummation* and *appetence* into the ethological literature to describe the *two* phases of *instinctual behaviour* (where an *instinct* isa mechanism consisting of a drive, a perceptual detector and a specific motor pattern). *Consummation* refers to the end phase or the final behaviour of, say, mating, eating, attacking, etc. This part of the instinctual behaviour is fixed and re-active in that it demands a specific (releasing) situation. But many organisms also apparently *search* forthese situations, and that is the *appetence* phase - a phase with more variability in applied behavioural patterns, and where higher animals can use their experience and even imaginative capacities (like planning capacities). This phase, it is assumed, corresponds to some *internal* state(s) in the organism. Biological functionalism puts an emphasis on *purposiveness,* or *goa/-directedness,* in organisms. But it is not just left at this. The investigator also attempts to find out about the *underlying elements and mechanisms* involved. [14]

This is the setting in which I want to situate an evolving planning capacity. There *was* already goal-directedness, spontaneity and subjectivity in the behaviour of biological organisms *before* there was any *planning.* These properties may be prerequisites for, or constitute a background for, planning but they as such do not require planning.

3. Goals, Problems and Interests

It should be clear that linked to the concept of *adaptation* is that of a *goal.* A form of behaviour control is biologically *adaptive* if it, in the end, increases the inclusive fitness of an organism (or of a species). In a *technical sense,* this is also the organism's utmost *goal,* i.e. to maximize its inclusive fitness.The organism is, again in a technical sense, *interested* in <loing so or has this *interest.* ('Interest' and 'goal' are roughly synonymous.) When a goal is not satisfied, there isa *problem* that the organism is interested to engage in.

Departing from this, it is, I believe, possible to construct a *biological goal - interest or problem - hierarchy* according to figure 1 below. (Cf. Bischof, 1985, p.331.)

---

14 In doing this, one departs from the earlier vitalist tradition which rather takes the recognition of purposiveness as a 'final explanation' of living creatures; as a 'final answer' to the question about 'the essence of living creatures'.

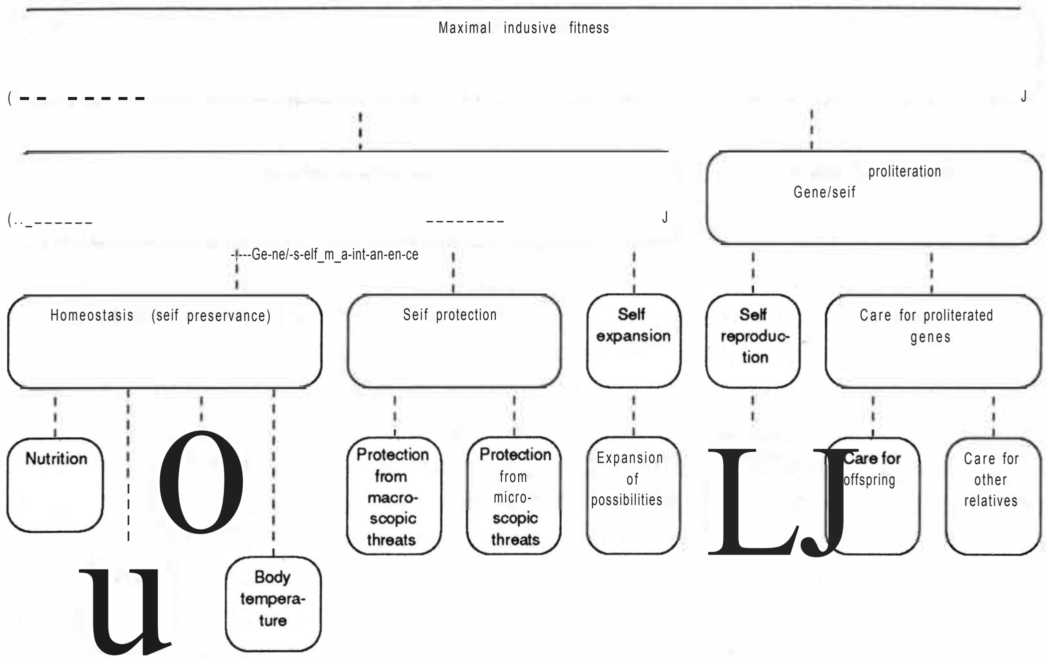


Figure 1. Biological interest hierarchy

In order to have one's genes propagate, there has to be some form of self*propagation* or *reproduction.* And in order to propagate, one must *maintain* one seif, that is, survive and maintain a state where one is capable of reproducing. These two interests - those of self-reproduction and of self-maintanence- make up the second level of the biological interest hierarchy (see figure). They are goals or interests *by definition* and are the same for *all* (biological) *speeies.* But these interests in turn *generate* further goals and interests. 1f survival and reproduction are enhanced by tendencies of preserving certain states, or to now and then reaching certain states, further interests will, namley, be generated: such as interests in mating, in getting nutrition, in regulating one's body temperature, in avoiding enemies and corresponding goals like having a supply of nutritional substances in the body, having a certain body temperature, being in a state where there are no threats around, and so on. The two interests at the highest level are the same for any species. On lower levels there is more variability. For instance, an organism *may* or may not have an interest in *sexual* reproduction (it may have an interest in self-fertilization instead). It *may* or may not have an interest in caring for offspring etc. The hierarchy depicted above, however, does, I believe, represent the basic biological interests and problem domains of all *mammals,* but to extend it to describe more detailed and specific interests and goals, there will be variation relative to the constitution of the *particular species* in its *particular environment* (the ecological niche it normally occupies and the more contigent conditions). There are specific interests such as interests in detecting mates in *particular ways* (say odor detection), interests in impressing mates in particular ways, interests in detecting and getting hold of food in particular ways, etc. (Cf. Dennett, 1984, pp.21-22.)

It is important to realize that the hierarchical structure of interests, goals and problems implies that the behaviour of an organism can always be described at several levels. For instance, we can say of an organism that it is at a particular moment engaged in all of the following: detecting a branch that can be a tool for termite-fishing, searching for a tool for obtaining food, getting itself something to eat and securing its survival. Or, to take another example, we can say that an organism is engaged in licking a wound, avoiding an infection, maintaining its health, or preserving itself.

A *subgoal* is a state, the reaching of which - given certain circumstances - contributes to or is a precondition to the reaching of a given goal. Thus, getting a tool for termite-fishing can be a subgoal in relation to the goal of getting something to eat, and staying clean can be a subgoal in relation to the goal of maintaining one's health. I prefer to use the notion of subgoal in a relatively wide sense, covering some different variations, as I will illustrate with some examples:

• Getting rid of pain can be a subgoal of the goal to keep intact and healthy. Obtaining this may (but does not necessarily) contribute to a maintanence or obtaining of the goal, as it merely works as a kind of signal.

• Getting some paint and brushes can be a subgoal of the goal of getting one's walls painted. The reaching of this state may be a precondition for and 'be on the way to' reaching the goal.

• Going to a restaurant or cooking at home can be subgoals of the goal of eating. These can be 'on the way towards the goal' and can also be seen as two instantiations or specifications of a precondition for eating, namely, to have something to eat.

The goal that - by an agent or an observer - is *treated as* the most general or the final goal for a behaviour in a given situation I call *the overall goal.* The agent's overall goal is the most general goal that the organism is actually sensitive to, in the sense that when it reaches *this state,* then the organism terminates its behaviour. The overall goal from the point of view of an observer is what she considers to be the highest goal that is relevant for a particular description or explanation of the behaviour.

Note that I use the terms 'goal', 'problem' and 'interest' in a *technical sense.* They cover *more* than what we normally understand by goals, problems, and interests, in that it is *not required* that an organisms *knows* of its goals, interests or problems. In this technical sense an organism can have goals, problems and interests although it has no inkling of them. A *problem* does not *have* to be perceived, represented or known by the organism, a *goal* does not have to be set by it, etc. [15]

There is another respect in which the notion of *goal* presented here is more general than the typical use of the notion. It covers what one may call *goals of achievement* (to have getting hold of that piece of food now as a goal) as well as *goals of maintanence* (to have the maintenance of one's body temperature within a certain range or to stay healthy as goals) as well as *goals of expansion* (to have getting hetter at running or improving one's wealth as goals).

The *basic mode/* of a goal in my sense is that of a 'desired state' or a 'desired value' of some quantity or parameter. This 'reference'- what is 'desired' - can be a particular value or a range. A goal of maintenance is one where it is desired to keep this reference state or

[15] For instance, there are no organisms perceiving a problem of proliferating as many of its genes as possible in coming generations and setting up a corresponding goal.

value continously stable. A goal of achievement is one where the state or value is desired only during a period of time. A goal of expansion, finally, is one where what is desired is not properly a value or spectrum but a *change.* A goal of expansion is in a sense open-ended, and in this sense it may never, at any moment, be' completely satisfied' even if an increase at a given moment means a relative satisfaction. In this respect, goals of expansion differ from the other two kinds of goals. [16] It should be noted that the most fundamental goal of the biological goal hierarchy - the goal of maximizing ones inclusive fitness - is a goal of expansion, as well as the goal of self expansion.

Apart from the hierarchical ordering, there is another difference in the status among the biological goals or interests represented in the biological interest hierarchy. We can see this by starting with the following consideration. It is evident, from an evolutionary perspective, that an organism should in the first place engage in immediate survival (dealing with nutrition, maintaining body temperature, keeping enemies away, etc.) and in reproducing. I call these *primary* or *serious* interests. But, if there is 'time over', bow ought the organism, from the point of view of evolution, use this time? This, I propose, is where the interest of *seif expansion* (see figure 1) enters. At the bottom, this goal concerns the expansion of one's *chances* or *capacities* to deal with serious or primary interests and problems. The prime activities that relate to the interest of self expansion are *exploration* and *play,* which involve learning new possibilities, gaining competence and exercise. Offensive attacking is also included. The interest in self expansion, you could say, is the interest in expanding- one's territories in a literal as well as in a figurative sense.

It is apparent, however, that the kind of activities I just mentioned, may involve danger. They are not undiscriminately advantageous, and so it seems desirable to balance such behaviour with cautious behaviour and withdrawal. My suggestion for how this is obtained- which builds upon ideas of Bischof (1985) - is the following. We posit an *interest* in 'having something novel to deal with; perceiving something novel, doing something novel'. You may think of this as a parameter that, at least in some respects, is parallel to, say, the parameter of body temperature. If the value of the parameter is too low (there is not enough novelty), it can be increased by activities such as going on an exploration tour, playing around, attacking, etc., and if it is too high, so that there is too much new information and novelty, the organism reacts by being cautious, withdrawing, retreating, or resting. There is thus *both* an interest in having something novel to perceive and in doing something novel, *and* an interest in 'familiarity' of what one perceives and does. (Cf. Bischof, 1985, pp.240-248.)

Now, the interest in self expansion and the interests derived from this; namely, an interest in novelty and an interest in familiarity, are not *primary* but *secondary* interests. This simply means that an organism will not engage in behaviour that relates to secondary interests as long as there is a current primary interest. If it has a problem conceming nutrition, threat, or cold, and so on, it will not engage in exploration, play, resting, withdrawal, or offensive attacking. (Cf. Lorenz, 1973, p.188.) Yet, even though the interests related to self expansion in this sense are *non-serious* or *secondary,* I want to emphasize

[16] I am not arguing that these are categories with sharp boundaries; goals can be construed in different ways, in particular as they coexist with and are influenced by other goals.

that they are *fundamental biological interests* and that the corresponding activities, such as exploration and play, are important biological activities, just as eating, mating and defensive behaviour. I think it is fair to say that the recognition of this is another feature of the biological functionalist tradition. And as I hope will become clear, this recognition, as well as the other ideas I have proposed in this section, does play a role for my analysis of planning.

4. The Concept of Motivation

As I have presiously stated, the terms 'interest', 'goal', and 'problem' are technical terms that cover more than what is *usually* understood by these nations. But how then are *such* interests, goals and problems to be characterized? To delineate goals, interests and problems that *are known of* by a system (an organism, a subject), I will use the concept of *motivation.*

In biological systems one finds mechanisms that continually *check* certain *internal states of the system* - plus *environmental states relating* to these. The states checked are states that are important for the organism's well-being and survival; states that relate to biological goals (of maintanence, achievement and expansion). The checking mechanisms detect *discrepances* between *current* and *desired states.* In other words, they detect *problems* or *current interests.* The occurrence of such a discrepance is then reported to what I call *the motivation-system.* This system has an evaluative and coordinative function. It assesses the insistance and urgency, and so on, of the various discrepances and compares them. On the basis of these evaluations, the system's resources (for behavioural as well as for internal activities) are directed towards one or some particular interests and problems.

Now, a *motivation,* in my usage of the term, is *aperception* of such a discrepance where the perception *can direct the resources of the system by initiating specific activities* in order to eliminate or reduce the discrepance. *Prototypical* motivations are those that relate to basic bodily interests; such as hunger, thirst, sexual arousal, pain, cold, fatigue, and so on. But there are also basic motivations that are less apparently bodily or physical; such as curiosity, anger and fear. And furthermore, there are more and less basic *social* motivations, like loneliness, desire for vengeance, envy, guilt, shame, and so on.

In the same way as there are primary and secondary interests, there are *primary and seconda,y motivations.* The secondary motivations are simply those that make organisms engage in secondary interests. There are the (secondary) motivations of boredom and curiosity (as perceived discrepances) that motivate playing and exploring. And on the other hand, in situations where there is *too* much novelty and unfamiliarity, there are motivations such as anxiety and wish for withdrawal that take over. Sometimes, however, *too* much novelty can be perceived as 'unfamiliar and scaring but challenging' and motivate exploration, which may involve that the organism becomes more familiar with, and gains some control of, the novel and unfamiliar situation/object/event. (Cf. Bischof, 1985, pp.240-248.)

In sum, the *motivation underlying a certain behaviour* or activity is *the perceived discrepancy* that *'moves'* the organsism in a particular way: that makes it eat or search for food right now instead of doing something else, that makes it flee from a situation instead of exploring it, or that makes it search company right now instead of staying alone. Furthermore, motivational mechanisms play an important role in *reinforcing* activities and behaviour. Typically, the reduction of the perceived discrepance, such as the reduction of hunger, tiredness, cold, anxiety, etc., is accompanied by some *pleasant* experience (of satisfaction, comfort, etc.). This occurs either in connection with consummatory behaviour or in connection with the arrival at some goal situation.

The core of the issue ofmotivati011 is this: It is the fundamental condition of a living organism with a certain repertoire of behaviour that *at every momenta choice must be made: some or other* behaviour pattem, some or other movement, some or other activity must be selected.

An organism may have more than one current *interest* and problem at a time. It can, say, simultaneously have too little nutrition, a temperature that is too low and be under threat. However, in general it is not possible to deal with more than one interest at a time. For instance, one cannot simultaneously eat, seek a shelter and defend oneself against an enemy. The simplest solution to this, which seems to be implemented in basic motivation-systems in organisms (possibly in the motivation-systems of all organisms except humans) is to always let only *one* interest *dominate* and actually take over; that is, to let one interest be the only one that gets through or dominates *as* a *motivation.* For instance, to let the organism be hungry and nothing but hungry even if there is also a *physical need* for sleep or for warmth; to inhibit tiredness and freezing and deal only with the interest in food and first thereafter with another interest. That is, at any time, the organism concentrates fully and solely on *one* interest or problem. *(Which* interest that dominates depends on the motivation-system's evaluation of the insistance, urgency, etc., of the various discrepances - which in turn depends upon factors internal as well as external to the system.) I call this *'the principle of one-motivation-at-a-time'.* (Cf. Bischof, 1985 p.255, pp.291-2.)

An organism with such a motivational mechanism will never be 'driven in two directions'. Conflicts between interests are solved at an earlier stage. Thus there will be only one current motivation, even if there may be several current interests. This seems to be not only a simple butan obvious *basic* solution. Note that this solution shall work in organisms that are *not* capable of *rejlecting* over their problems, goals and actions, of thinking through actions, etc., and in this way decide what goal to pursue; whether to eat or sleep, whether to look for food or fora sleeping place, whether to eat or drink. But these 'decisions' (basically decisions concerning resource distribution in time) must be made by a motivation-system at a *non-rejlective leve/.*

Toere is ethological evidence for the principle of one-motivation-at-a-time as a basic solution. It seems to work in this way for most organisms and in most situations. There *are anomalous cases* where animals seem to experience motivational conflicts, but these are rare and can, to my mind, be taken as illustrations of what happens if the mechanism is put out of order.

I will refer extensively to the principle of one-motivation-at-a-time during this thesis. To be more precise, however, it ought to be specified as 'the principle of *one-primary-motivation-at-a-time*'. The reason is the following. Primary interests have priority over secondary interests. Thus primary motivations have priority over secondary motivations. For instance, hunger has priority over curiosity, fear of a <langer has priority over a wish to rest. Playing, resting, 'just exploring for no reason', etc., is only undertaken in the absence of a primary or serious motivation. However, even though an organism will thus

not stop searching for food or escaping from an enemy in order to start playing or resting, it is possible to have secondary motivations *infiuence* behaviour that deals with serious interests. Anxiety, for instance, may make an organism cautious in its search for food, influencing where it goes, where it searches, etc., whereas curiosity may influence an organism, say, so that it will stop and begin to explore something that looks edible. In general, secondary interests may be combined with primary interest in different ways. And so, when I indicate the motivational state of an individual, I sometimes state only one motivation (primary or secondary) but sometimes - when this is relevant - I give a primary *as well as* a secondary motivation. Ishall in the following use this notation to describe motivational states: 'M(primary interest; secondary interest).

Considering the interest hierarchy on page 24, one realizes that there do not exist motivations in my sense - that is, perceptions of states where these perceptions can direct specific behaviour in order to change the perceived state - for any interest or goal. There *are* motivations for the interests of sleeping, getting food, mating, avoiding enemies, maintaining body temperature, eating, drinking, staying clean, hiding, fleeing, defending oneself, exploring, etc. But moving upwards in the hierarchy there is no *motivation* for the interest of self-maintanence or for the interest of maximizing one's inclusive fitness. There are no specific discrepances that can be detected and that can initiate some particular behaviour to eliminate the discrepance. Also, if one considers what happens on the levels of neural and biochemical organization, one can speak of goals, interests and problems but not of motivations. There are of course mechanisms that detect what happens here, and these mechanims deliver basic information to the motivation-system. But these states and processes are not perceived by the subject.17

In sum, motivation concerns *currently perceived* discrepances from desired conditions, (discrepances between a cun-ent anda desired cun-ent state). And I suggest that what we understand by feelings, drives, emotions, moods, etc., are different kinds of motivations. Motivations are the basis for spontaneous behaviour - appetence or search behaviour: I am thirsty and will not merely sit and wait for something to drink to appear (that I can react to); I am tired and will not merely sit and wait fora place to sleep to appear; I am lonely and will not just sit and wait for some company to come. Instead, I *spontaneously* move; I search.

But now we come to a question that will receive considerable attention in this thesis. What about a system that can also deal with *non-current* potential problems and *potential motivations;* like dealing with the problem of finding a place to sleep even when not currently tired, or dealing with future predicted hunger? It is evident that human beings can do this. For instance, before going on a bicycle tour, you may devote time and energy to thinking through and making preparations so that you will have something to eat and drink and somewhere to sleep during your tour, although you are not the least hungry, thirsty, or tired at the moment and have no current motivation to eat, drink or sleep. Or

17Tois becomes apparent in cases of anomaly where the normal causal connections are disturbed, like say when the physical or bio-chemical nutritional need is satisfied but the organism is still hungry and continues eating or the other way round. Such effects can be 'obtained' by manipulating certain areas of the brain in animals. Similar effects can be found for temperature regulation. (But during normal circumstances the motivations are connected to the neurological and biological situations, and the behaviour regulated by these motivations will lead to the realization of the uppermost goals as well.)

you may think about what to do in the case of some potential threat, even though you are not threatened at the moment. (Cf. Anita in the first example, pp.3-4.) That is, the distribution of resources in human beings - what they do and think of and attend to - is also directed by nonactual, only potential interests. They may think of problems and actions where the overall goals do not relate to any current, but rather to just potential interests and motivations. Furthermore, humans may consider and act in relation to goals for which there *is no* motivation in the basic sense, like an individual who thinks of having a lot of children with the goal of securing his future or even with the goal of proliferating his genes (for which there is, we have agreed, no motivation). It is not excluded however that his actions are *also* driven by a motivation for mating. In general, the issue of proper motivations is complex in humans. Say, for instance, that a person is on a diet and so at a given moment refrains from eating although he is hungry. It is not evident that there is a perception of a discrepance that makes the person act in this way in order to counteract the discrepance, in the way that hunger may motivate one to eat. Yet it is *possible* that some social motivation like a feeling of inferiority or a desire for self-assertance - which may be construed as such a perception of a discrepance - actually influences the person's acting.

Finally, not only behaviour but also *internat activities* must be motivated in some way or another. Different intemal activities can compete with one another, and internal activities can compete with behavioural performance. *Planning* must be *motivated,* and there must also be a motivation for dealing with particular problems and goals when planning. These are issues that I will address in this thesis.

5. The Relevance of Biological Functionalism for my Investigation of Planning

In a substantial way, biological functionalism supplies the rationale as well as the inspiration for a main part of my work. It is this framework that supports my attempt to regard the phenomenon of the planning of action as a *biological* phenomenon, and not as something that should primarily or only be investigated at the level of psychology or sociology. I rely upon the general idea of the *human mind* as at least partly shaped by evolution and the belief that one can understand it hetter by investigating its evolutionary origins. Briefly, the idea is that human planning of action - like planning a dinner or planning a journey or planning one's aftemoon - has its roots in simpler forms of planning - like a chimpanzee's planning to get some food or planning to get over an obstacle - which, in turn, has its roots in more primary cognitive and behaviour-regulating functions. I do not regard human planning as an independent kind of behaviour-regulating function with its own new substrate and sub-functions but rather as a function that builds upon and is integrated into more fundamental forms of behaviour control. Instead I regard it as a high-level intellectual feat that is only possible in virtue of our being endowed with a set of simpler, more basic competences that solve more fundamental survival problems.

The overall structure of the thesis bases upon the biological functionalistic approach that I presented on page 20. Part Il and Part Ill aim at a systems understanding and an evolutionary understanding, respectively. Furthermore, as will become clear, I shall use of many of the ideas and concepts that I have introduced here conceming goals, interests, motivations, appetence behaviour, adaptiveness, and so on.

4. ON THE SYNTHESIS OF COGNITIVE SCIENCE AND BIOLOGICAL FUNCTIONALISM

1. Introduction

I now want to explain why I have chosen cognitive science and biological functionalism as frameworks for an investigation of the phenomenon of the planning of action. First of all, I believe in the independent value of these two approaches for dealing with issues in the domain of psychology and behaviour control. However, I also believe that they, in important respects, *complement* each other and that therefore the synthesis of these approaches is suitable for my investigation of planning.

In this chapter I will show bow *system theory* and *cognitive constructivism* are meeting points for the two frameworks, indicating bow a biologically grounded *system theoretical approach* can be helpful to a cognitive science approach and how cognitive science and biological functionalism mutually support each other as concems the idea of *cognitive constructivism.* The section on *cognition and emotion* shows bow a biological approach modifies mainstream cognitive science. Here I also situate my framework as non-behaviouristic. I then conclude the chapter with a section in which I discuss the relation of my investigation to other kinds of research on the planning of action.

2. The Systems Approach

There is indeed a 'discipline' or approach that traditionally bridgesa *biological* [18] study of the mental life and behaviour of organisms on the one hand, anda study based upon *information-processing ideas* on the other hand. This is *cybernetics* or *system theory.*

A *system* is a class of elements that are connected in such ways that they have a comparatively immediate or direct influence on each other. If one does something to one element this will influence the other elements, i.e. the rest of the system. The elementsofa system stand in closer relationship to one another than to what is outside, i.e. the system's environment. 'System' isa hierarchical notion; the elementsofa system can be systems themselves. A system can be composed of *sub-systems.* The set of *relations* between the elements or sub-systems constitutes the *structure* of the system.

System theory or *cybernetics* is the general study of systems in this sense. To use system theory to account for the behaviour of a system, or rather for *some* behaviour or some property it has, is to try to describe a structure of sub-systems and bow these interact in order to produce this behaviour or property of the whole system. That is, one tries to describe *the organization behind* the *observable* behaviour or output.

Presented below are some simple system-theoretical descriptions. Figure 2a isa general or abstract illustration of a feedback-situation. Figure 2b describes the behaviour of a system that moves in relation toa particular object-where the movements can be described as being aimed at the maintanence of a certain distance to this particular object. When

18 Taking biological in a wide sense as including neuroscience as well as ethology.

the object approaches, the system moves away from it; when the objects moves away, the system moves after it. And figure 2c illustrates an explanation of a system's withdrawal and exploring behaviour when it encounters novel objects. We find here a sub-system that is a perceptual evaluator of the degree of familiarity and novelty of objects, as well as a sub-system that regulates the system's tolerance and desire for novelty. When the current degree of novelty is higher than the desired degree or value (see figure) the system gets anxious and reacts with withdrawal when encountering a novel object, but if it is not, it gets curious and begins exploring the object.

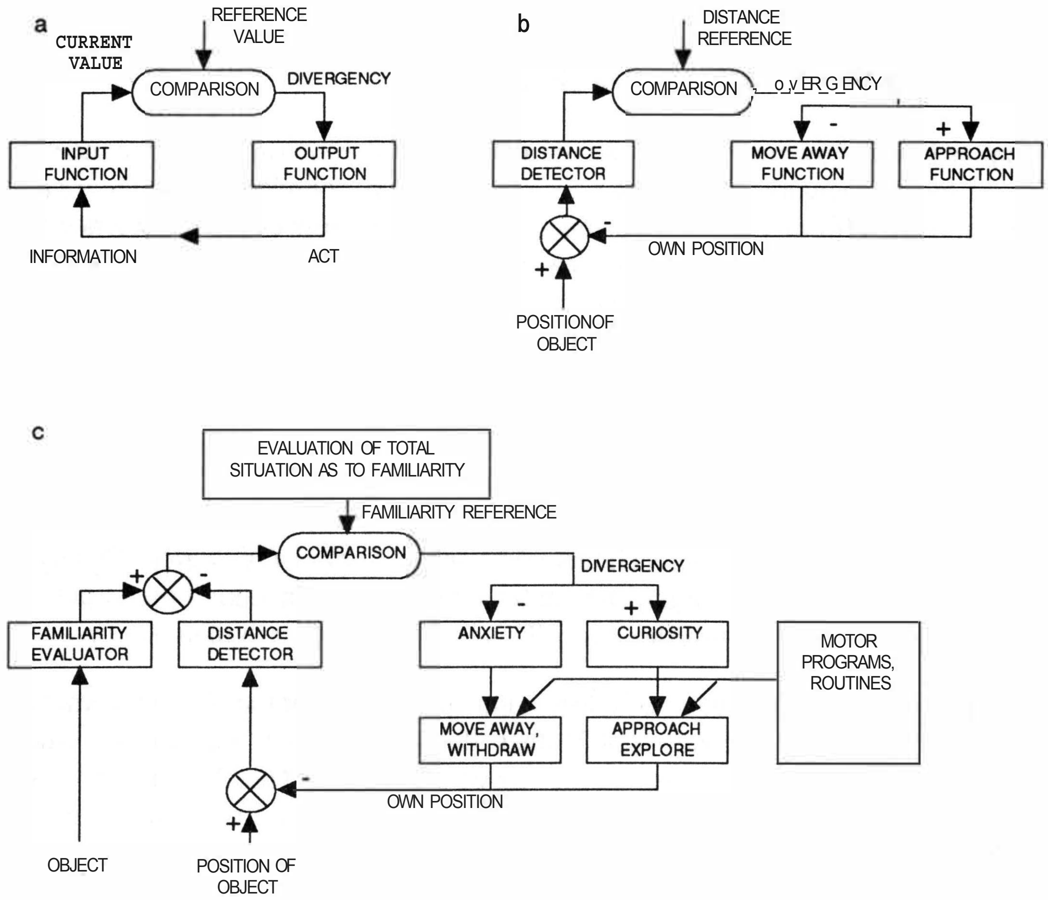


Figure 2. System theoretical illustrations

Observe that one is *not* concemed with just describing the organization of the *observab/e* output or behaviour of the system but the organization *underlying* such output. Thus, the use of a system theoretical approach presupposes a willingness to talk about the non-observable, and also a certain willingness to 'carve things up'. For instance in the example of exploring and withdrawal, we postulate elements such as curiosity, anxiety, a perceptual evaluator for novelty, etc., where none of these can be observed. The choice of a *leve/ of analysis* is important in a system-theoretical study: you must stop at some point, namely, at the smallest sub-systems, the internal structure of which one does not go into. For instance, in the examples above, everything concerning the *structure and realization* of the perceptual distance detector or of curiosity is left open. System theory is an approach that involves the construction of abstract descriptions. It deals with relations between *functions* and with flow of *information,* not with *substrates* and *energy* flow.

In addition to providing a format of description - ways of notation - system theory supplies certain concepts and ideas that I will describe briefly. A *reference state* or *reference value* is a state or a value of a certain *parameter* that is *'desired'* in the sense that *disturbances* in one or another direction are *counteracted.* The system achieves or maintains this state or value through *actions* (on the environment) that counteract disturbances. We say then that the system *controls* these parameters or quantities. The presence of *feedback* means that there is some causal chain or loop that relates the system to its environment so that the *effects* that the system's actions has on the environment are 'fed back' into the system. The system is *informed* of this, and the information that is thus fed back has effects on subseqent actions. (In *positive* feedback a disturbance produces an effect on subsequent behaviour that reinforces the effect of the disturbance. In *negative* feedback the effect is to counteract the discrepance. In other words, occurrence of negative feedback means occurence of control.) (See Sachse, 1971, pp.64-98; Powers, pp.285-286.)

In my opinion, a system theoretical approach is suitable for dealing with behavioural and internal activities of living organisms. It provides useful ways of thinking and a notation where the complex dependencies among sub-units, hierarchical structures and nestings that seem abondant in living organisms can be expressed. Thus I believe that it is useful for investigations within cognitive science. In effect, there are two problematic tendencies in many attempts to model human cognition and behaviour in cognitive science, both of which one could possibly come togrips with by means of a more thoroughly *biologically grounded system theoretical* way of thinking.

First, there is the practice of modelling a *restricted* aspect of human performance, to cut this loose and deal with it as an independent competency, in other words as a separate sub-system within the whole system (the human being or her mind). But from a biological perspective, aiming at consistency with an evolutionary picture of the emergence of mind, one might be more careful as to 'how one carves up things'. In particular, one ought to consider the idea that many 'high level' acitivities and achievements in humans emerge from the combined action of several lower level capacities which are more directly geared to the basic needs of a developing organism. (See Clark, 1985.)

Second, there is the practice of modelling an achievement, a competence, a mental operation, etc., in a very restricted domain or environment, and in this way simplify the problem that the system has to solve. Such a restricted environment is often spoken of as a *microworld.* Now, biological thinking *does* concede to the idea of restricted environments - ecological niches and also the 'Umwelte' (see p.22.) of different species. But the inspiration from the biological approach should then rather lead to attempts to deal with *whole* niches and the Umwelte of simpler creatures instead of concentrating on *fragments* of the human world (like the domain of chess-playing or the domain of restaurants). (See Clark, 1985.)

My use of system theory - of ideas as well as concepts and notations - will be extensive. I am interested in the *internal organization* of *systems* that are capable of perceiving, acting and of planning their actions. I also try to avoid the two tendencies noted above. I do not approach the planning of action as a separate high level function. I do not discuss a *planning-system* (a system that does nothing but plan) but a *planning system* (a system capable of more than planning). And I try to avoid the microworld problem by dealing with a whole environment and several capacities and activities of some *simpler* creatures.

3. Cognitive Constructivismfrom a Phylogenetical Perspective

In this section I will point to an area where it is evident that biological functionalism and cognitive science meet and complement each other. This is the area of cognitive constructivism. On the one hand, I will relate two objections to the idea of cognitive constructivism, which we know from chapter 2 is important in cognitive science, and show how this criticism can be met by taking a biological - evolutionary - perspective. On the other hand, I will show how some ideas of biological functionalism are elaborated and supported by cognitive science research. (That is, I indicate how both frameworks complement each other in this domain.)

The first of the two objections against cognitive constructivism is directed against the view that *allperception* is an issue of construction. The cognitive constructivist's claim is not only that *'high leve!* perception' is constructive; for instance that a specific object, say a sailboat, may *mean* different things to different perceiving subjects because of their differing experiences with sailing boats, their different knowledge of boats, their different moods, and so on. But there is also the claim that to perceive *objects at all* is to *construct objects. All* perception of and knowledge of objects is the result of constructional processes. There is no 'direct perception' as, for instance, Gibson (1979) has it.[19] Yet, presumably *all human beings*- as well as *all* individuals of certain other species-do *perceive objects.* Hence the objection: how can one be justified in talking of *construction* when there is no variability, no degrees of freedom? The answer to this, from a biological - evolutionary - perspective is this. There *are* degrees of freedom since phylogenetically there is more than one way of processing certain information. Nature has degrees of freedom in constructing nervous systems and minds; in constructing ways to select and produce representations.[20,21]

The second objection - which on the surface appears to be an objection from an evolutionary perspective - is the following. According to cognitive constructivism it is the *organism* that constructs and in this sense *selects its environment* (the environment it perceives, behaves in, etc.) rather than *vice versa,* namely that the *organism* is *selected by the environment,* as Neo-Darwinism has it. Therefore constructivism is not in line with evolutionary theory. This, in my mind, is a shallow and, in spite of its appearance, unbiological kind of objection. It is not the existence or lack of superficial structural parallels that decides whether a view is in concordance with evolutionary thinking or not. In

---

[19]Gibson claims that in the 'ambient light' there is objectively specified information, *directly* specifying what is there in the world. *This* information is *'picked up',* and there is no need for any mental supplementation, compensation or construction.

[20]As to 'higher cognitive processes' - processes involved in thinking, reasoning, planning, etc. - few want to deny that there indeed are degrees of freedom, and that these processes are *constructive.* And so, *constructivistic* theories of perception have less difficulties in accounting for the *re/ationship* between perception and these higher capacities - that is, to phenomena such as thinking, planning, memory, etc. - and in accounting for facts of interplay. If, as Gibson argues, we are to do without mental representation in experiencing and perceiving the world, we are left with a gap between perception and experience on the one side and the activities of thinking, remembering and making plans, etc., on the other side, as these latter processes clearly do involve operations upon mental representations.

[21]For a lucid discussion on the evolution of *object representation* see Von Glaserfeldt (1976).

somewhat more detail, the relations between behaviour, intemal environment and external environment can be described as follows. The organism, true enough, does select its environment in the sense that what is 'out there' is filtered through the organism's internal model in order to enter behaviour control. And in that sense it is the *internal mode/ or environment* that influences or *selects the organism's behaviour.* However, the internal model has itself gone through a long selection process, where the external environment is a crucial element. To put it somewhat bluntly: even if organisms *experience, act in, think of* and *plan actions* in their *internal environments* or mental models, they *survive* (or do not survive) in the *external environment.* Toere *may* well be *elements* in an internal environment that only have a very *indirect* relation to the external environment, ora misleading one, or none at all. Still there is reason to believe that many other elements do in some sense correlate with external elements[22], as the internal environment and the *behavioural capacities,* in the development of a species as well as in the development of an individual, evolve *together.* (Cf., for instance, Campbell, 1966; Lorenz, 1973, pp.17-20; Riedl, 1979; Vollmer, 1975; Wuketits, 1981.) In sum, an *evolutionary* approach to perception and cognition makes sense of and renders plausible cognitive constructivism even in a radical form.

On the other hand, certain ideas from the tradition of biological functionalism have, I believe, gained more substance through the growth of cognitive science. The aspects of living organsims that this tradition focuses on and describes - appetence behaviour, subjectivity, species relativism in perception and cognition, etc. - can be brought out and illustrated in research and modelling within the cognitive science framework. The advantage of cognitive science in this respect is that it moves from descriptive towards explanatory levels, in particular, in using the notion of mental representation. Cognitive science studies have, I believe, in a *more explicit* way than biological functionalism *reinstated the concept of mind* and pointed to the *need* for considering internal structures and processes in explanations of organisms and their behaviour. And in doing so, the biological functionalist conception of subjects and their intemal constitution as worthy of scientific investigation has been strenghtened.

The phylogenetic perspective on cognitive constructivism and the lines of thought I have here related, play a role for my study; in particular for how I see the basis of a planning capacity in functions involved in perception, and also for giving some more substance to the notion of mental representation that I rely upon.

The constructivistic ideas in biological functionalism as well as in cognitive science contrast with a *behaviourist* approach to behaviour and mentality. I have two remarks on this. First, it is a behaviouristic premise that the only thing we need to study to understand behaviour is environmental stimulation; present environmental stimulation and the organism's history of such stimulation. We need not study anything *internal* to the organism. Cognitive science and biological functionalism, on the other hand, emphasize the role of internal structures. And biological functionalism, in particular, points to organisms' capacity for spontaneous behaviour where the *spring of action* is internal and not external to the system (for instance, appetence behaviour). The *general* description of be-

---

[22]Where 'correlation' means that if the represented - in the external environment - would have (bad) other properties than it actually has, the representation - in the internat environment - would also have (had) other properties.

haviour is not an event or state in the environment that as such is responsible for causing behaviour, but rather a *discrepance* between the organism's perception of an event or state and some 'desired' event or state causing behaviour. Secondly, the behaviourist is of course neither interested in *differences in internal organization* between different speeies or different individuals. But this is, as we have seen, an essential issue for the approach of biological functionalism. Ethological data on species-specificity in perception and behaviour are indeed difficult to handle within the behaviouristic paradigm. These behaviours and ways of perception do appear in the normal course of development *without any obvious learning* - whereas the behaviourist wants to have all perceptions (stimuli) as well as all behaviour (responses) as *initially equivalent.*

4. Cognition and Emotion

In this section I discuss another issue that is of importance for this thesis and where a biological approach complements cognitive science research, namely, that of the relation between cognition and emotion. Much *mainstream* cognitive science research focuses exclusively on 'paradigmatical knowledge-processing', that is, on processes involved in *paradigmatically cognitive* phenomena such as reasoning, perception, memory, problem salving (in a strict sense), etc. Now, there are certain *assumptions* that often go along with this research. There is the conception of the human being as a - or *the* - *rational creature.* Man has an intellect; he can *reason* about what he does or will or should do. This *new, separate* and *independent* capability, that we describe in terms of rationality, intelligence, capacity for reasoning, it is assumed, *emerged* in humans and *characterizes* the species and its behaviour. And in order to study and explain *human behaviour* we should thus concentrate on the capabilities that underlie human perception, reasoning and thinking. It is these capabilities that *primarily* guide human behaviour. There are of course also certain *irrational elements* such as emotions, drives, feelings, moods, etc., but these are to be seen as disturbances, a kind of noise.

Thus, according to mainstream cognitive science, to understand human behaviour, one ought primarily to study *cognition,* not *emotion.* And also to *understand cognition,* of course, there is no need for dealing with emotions. That is, what we find is an acceptance of the *classical subdivision* in psychology of *cognition* and *emotion* as two separate domains. This division and the conception of emotional reactions as *irrational,* of emotions as something that disturbs intelligent, sound actions, goes far back in the tradition of Anglo-Saxon psychology.

From a biological perspective, however, the inadequacy of these assumptions is quite evident. First, it is misleading to relate a distinction 'cognition-emotion' to a distinction 'rational-irrational'. 'Irrational' has connotations of 'not intelligible', 'disturbing', 'not adaptive', 'not explicable', etc. But here the biological tradition departs from traditional (anglosaxian) psychology, in that it asks for the *adaptiveness* of emotive as well as of cognitive elements, which are all seen as evolutionary products (cf. chapter 3). To distinguish *explicit* thinking and reasoning from other kinds of 'intelligent' or adaptive kinds of behaviour-regulation, it is in my view more adequate to use terms such as 'reflection', 'reasoning' or 'contemplation' than that of 'rationality'.

Furthermore, assume that it is the case - which is likely - that emotive elements are parts of phylogenetically older mechanisms for controlling adaptive behaviour. Given the way evolution works, the emergence of a new wholly separate and independent form of behaviour governing- 'reason', 'thought', 'cognition', 'the intellect' - is not likely. It is more plausible that there would emerge certain new functions that would *build upon, extend* and become integrated with the older *'ratiomorphous'*[23] - *rational* but not *reflective* - functions, to make up new and more complex system functions. The older functions would neither just *disappear* nor be left as a kind of *disturbance.* And so, in sum, to *understand behaviour and cognition,* these older emotional functions should also be taken into account. The traditional separation of cognition and emotion as independent domains that have not much to do with one another is inadequate.[24]

These considerations play an important role for my approach to planning. I do not regard planning as a function that belongs to an independent sphere of reasoning and reflection, separated from what we understand by emotions, feelings, drives, instincts and so on. I do not think that once organisms are endowed with a planning capability, they leave behind more primitive or fundamental behaviour goveming mechanisms. Instead I treat the capacity for planning as a kind of additional capacity that modifies but does not replace more fundamental mechanisms of behaviour control. And I consider motivational and emotional aspects of a planning system as well as more purely cognitive aspects.

5. The Relationship of this Project to Some Other Research on the Planning of Action

The background for this study, I have said, is not made up by any specific studies on the planning of action but rather from more general research traditions that deal with mentality and behaviour. In this sense, that is, since it does not build upon any other studies of planning or any main research traditions dealing with planning, it is a solitary project.

There *is* a great deal of research, mainly in psychology and in Artificial Intelligence, on planning in general as well as on human planning. But, as I will try to make clear, a main part of this kind of research is based upon certain assumptions that I do not share. Starting with AI, there is a main research tradition, that might be called the classical planning approach. (For instance, Miller, Galanter, and Pribram (1960), Sacerdoti (1977), Wilensky (1983), Hayes-Roth and Hayes-Roth (1979)). The task in AI is often to construct a 'pure planning-system', a system that 'cannot do anything but plan'. That is, one aims at producing programs or systems capable of *planning* but completely incapable of *acting, perceiving* and *interacting with an environment.* Thus, a *capability to plan* is the *only* capability *desired* in the system.

Furthermore, one often deals with well-defined planning tasks - likethat of performing certain errands duringa certain time span or salving the Tower of Hanoi or some other intellectual puzzle. The goals and premises are explicitly given to the planner, and the 'essential information' is specified. And so, as soon as the problem is posed the system shall start constructing a plan of action. These studies, thus, focus on what occurs in the mind of the planner from the time a problem is posed until it is solved or abandoned.

23This term was coined by Egon Brunswik to refer to processes that are functionally and formally analogous to conscious reasoning but are not otherwise related to it (see Lorenz, 1973, p.155.).

24Cf. Sloman, 1987, who, from a partly different perspective, reaches similar conclusions. Cf. also Pugh, 1977.

Issues concerning how a problem is *found,* how goals are chosen, how to decide *whether* or not start to plan and *whether* or not to continue to plan, are neglected.The motivational side of planning is disregarded (that is, questions such as 'why shall the system plan now?', 'when is it desirable to plan?', 'how can we make the system plan on these occasions?' etc.). This is of course not unexpected. The system does not really have to be motivated for planning, that is, to chose planning before something else, as there is nothing else it can do (like performing a certain action, performing some advanced perceptual analysis, going to sleep, etc.)

Another reason underlying the idea that 'planning is always desirable' and underlying the neglect of the issues of motivation is the *conception of a plan* within the classical planning approach. This conception seems to be the following. A plan is like *a computer program,* from which action can result through a process of execution - and to plan is to construct such a program. *Plans* underlie all behaviour in the sense that all observable structure in behaviour comes from a plan with the same structure. A plan is thus the universal formula for behaviour production. It is, in general, plans that in the end specify and causally engender behaviour. Thus, if something goes wrong in behavoiur - i.e., if some behaviour is maladaptive - there must be something wrong with some plan. Possibly the *planning capacity* - which produces the plan - is not *poweiful* enough, and one ought to try to improve it; for instance, by giving the planner more memory, a hetter representation of its world, more knowledge, hetter ways of gaining access to its knowledge, a capability to handle longer sequences of actions or a greater number of goal-means-relationships, etc. But it can - of course - *not* be the case that *plan production* and planning as a general solution is wrong.

In brief, it is assumed that:

(1) All that a planner shall do and needs to do is plan. (Planning is an obligatory activity.)

(2) Planning of action is inherently desirable; that is, in general, planning one's action is desirable and 'a good thing'. (Cf. Goodnow, 1987.)

These assumptions are shared with much research on planning in psychology. In contexts that conem training and development of the capability of planning, there is an underlying attitude to planning 'as a good thing'. And the motivational side of planning seems until recently to have been neglected to a great extent.

In sum, the planning of action has traditionally been systematically studied only in isolation from other mental and behavioural capacities, mostly with techniques from classical AI and standard cognitive psychology. This, however, is not satisfactory if we want to understand the phenomenon of planning and its role from an evolutionary perspective, or if we want to construct realistically intelligent artificial systems that can do other things than plan; that is, systems that can and want to plan but also can and want to do other things.[25] In opposition to the assumptions above, I want to stress the following:

---

25 And the purpose of this study, as I have said, is *to* obtain a more realistic characterization and more complete conception of the phenomenon of planning, although this is achieved, to some extent, at the expense of being tentative and not so detailed.

(1) A biological system is *in the first place* a perceiving and behaving system. (And so I am not interested in discussing the design of a system that only plans actions but is not capable of perceiving and acting.) Planning in a biological system is *one* resource-demaning activity that competes with other activities. Planning can, for instance, neither be performed simultaneously with advanced perceptual analysis nor with the production of some advanced behaviour. Thus the motivational side of planning cannot be neglected.

(2) Planning is an activity, the adaptiveness of which varies in different situations. It is not always good to make plans for actions. True, our planning capacity must, as a biological function, have an evolutionary value, but this does not imply that planning is useful under all circumstances or that all aspects of the capability are so.

Recently, however, there has appeared in some AI eireles what I call 'the new wave', due to a dissatisfaction with the limitations of classical planning conceptions and techniques. Here one departs from the conception of a plan as the general formula for behaviour production as well as the attempts to understand and model planning and plans in this way. Instead *actions* are spoken of as being *situated;* meaning that actions take place in *particular* situations that are *difficult to specify in advance.* This is a reason why planning and plans will not be sufficient to drive an autonomous agent. Agents also need capacities to respond flexibly to the moment-to-moment contingencies of their environment which cannot be sufficiently specified beforehand in a plan.

In 'the new wave' there is an emphasis on the view of actions as *improvised* in contrast to planned. An important influence is L. Suchman (1978), who presents a conception of actions as *improvised within a social framework.* People *do* use plans. Prime examples are maps, recipes, etc.; that is, collective and extemally represented plans. But these plans, according to Suchman, do not *predetermine action patterns* and do not specify in advance how to proceed with some action. Instead they are used as *criteria* for judging what one has *just done* or *is doing during improvisation.* They are criteria that one may bump against, available to individuals for judging progress of action, while acting. In conclusion then, according to 'the new wave', it is *not* plans alone that are responsible for the structure of behaviour: a plan is not a general formula for behaviour production, and the idea of a plan as an intemal representation in an individual is not taken as the *paradigm of behaviour production.* And, consequently, plans and planning are not regarded as inherently desirable.

I am sympathetic to this shift in that I share the criticism of the classical. approach, that is, the objection to the idea of a plan as an intemal representation in an individual as the paradigm for behaviour production and to the conception of plans and planning as inherently desirable. Yet I find a tendency in 'the new wave' to move too far towards the other extreme, and deny or de-emphasize the phenomenon of *planning* in the sense of an individual's *producing* an intemal, mental representation of action. One seems to regard this either as a *nonexistent* phenomenon (an illusion) or as something that only plays a *marginal* role in structuring people's actions. The plans that exist and play som role are collective, social and extemally represented. This view indeed fits in with trends in anthropology that stress a view of *cognition in general* as a collective product, not bound to any individual minds. (See, for instance, Geertz, 1973.) I do not share these convictions. To the contrary, I do believe that people really *generate* mental representations of potential courses of actions and that these to a considerable extent structure their behaviour. It is this phenomenon, which I do not consider marginal and the role that it plays in human activity, that I seek to understand.

Among the conceptions and definitions of 'the new wave' in AI, one also finds a tendency to reserve planning capability for (highly) social creatures. 'The new wave' introduces conceptions of planning as the "intemalization of social communication about activity" (Agre and Chapman, 1987, p.272.) or of plans as "communication in natura! language" (Agre, 1990, p.1007.). I do not question that such conceptions are valuable for dealing with certain aspects of planning - I believe indeed that many aspects of human planning might be helpfully understood this way - but I do not believe they are adequate for a *general* treatment of the character and role of planning in humans and in other creatures. This is because I do not think that planning is intrinsically a social activity or phenomenon, nor that it depends upon communicative capacities. The basic phenomenon of individuals structuring their own future actions - thinking about what to do and bow to do something - is, I believe, more deeply rooted. In this way I oppose a tendency to view psychological phenomena in humans as completely and fundamentally grounded in 'advanced' cultural and social structures, requiring lingustic capabilities, and so on. I believe there are principles to learn about by considering more primitive forms and contexts of psychological phenomena.

In psychological research as well, there has been a shift conceming the two assumptions I mentioned on page 38. One is leaving behind the idea of planning as a mandatory activity and as inherently good and desirable, and motivational aspects of planning are taken into consideration. See for instance Nuttin (1984), or the anthology *Blueprintsfor Action* (1987), where several of the articles explicitly deal with the motivational aspects of planning. In general, researchers have started to take the various *costs* of planning into consideration. It is realized that in many situations it is better not to plan, either because planning will not lead to the agent more efficiently attaining its goal(s) (by gaining time or some other resources), or because of other critera that make planning undesirable, even when it is feasible and valuable in an objective sense. Thanks to certain more differentiated analyses of the concepts of planning there has also been a shift to leave behind the previously dominating conception of planning and plans as something extremely prevalent, that in principle underlies all structured behaviour. Again, I welcome these shifts. Yet, I would like to put them in a more general context. In psychology as well as in AI, namely, the discussions and the shifts seem to focus on the *social* aspects of planning, whereas I believe that questioning the two assumptions on page 38 is also necessary for dealing with planning in non-social creatures.

PART TWO: A SYSTEMS EXPLANATION OF PLANNING

"'WouCcf you teff me, p(ease, wfiicfi way I ougfit to wa(f(_from fiere?" "1liat depends agood deaC on wfiere you want to get to," said tfie Cat. "1 don't mucfi care wfiere - " said Afice. "'IJien it doesn't matter wfiicfi way you waCk," said tfie Cat.

Lewis Carro((, Afice in 'WonderCand

5. INTRODUCTORY ANALYSIS OF THE PLANNING OF ACTION

1. Introduction

My first aim in this chapter is to give a general definition of planning. I then attempt to distinguish planned behaviour from behaviour that is regulated in other ways and discuss criteria for deciding whether an action is planned. I describe the distinction hetween *immediate* and *anticipatory* planning. In a following section, I discuss biological planners, or planners in nature, and finally I present the illustrative example of the Berry-Creatures.

2. A General Definition of the Planning of Action

Before I present my proposal for a general definition of planning, let me give some *examples* of planning by describing some situations where an individual plans (some of) his actions.[1]

(1) A person is hungry and wants to have something to eat with her tea, but there is notbing in the pantry. She think:s about what to do: Go and buy something? The shop nearby is closed for the holiday. To go to the supermarket will take a long time. Then she think:s of those scones with nuts which can be quickly prepared. What is needed? Flour, margarine, salt, milk, nuts, raisins, baking powder... All is there. There are indeed these nuts left from Christmas. But of course then you need a nut-cracker. - Hm, go and buy one? No. How can you crack those nuts then?... She remembers than that she once saw a friend cracking walnuts with a garlic press. - Alright, she will tty that. And if it does not work, some of the neighbours ought to have a nut-cracker. Yes, she can ask Mrs Anderson preferably, and maybe Miss Preston... So, she'll turn the oven on and start.

This example presents a person who *thinks out a way* to get something to eat with her tea, which is her *goal.* She suggests to herself a sequence of actions that shall transform the current situation into this goal situation (which she quickly specifies as 'having some of those scones with nuts with her tea'). Furthermore, note that some of the actions that she suggests to herself are *discarded,* whereas others are assented to.

(2) A man is a tourist in a city. He is out shopping at the Main Square when there is a heavy rain shower. He gets soaked and is freezing cold. He wants to go back to his hotel at John's Park to take a warm bath and change clothes before mayhe going out again (and this time with an umbrella). He think:s: How do I get to the hotel then from here? I might just go back on Rose Street and turn left where it ends, yes I'd hetter do so. It means, however, walking in the wrong direction for a while. Is there a hetter way, a shorter way? Well, it may be shorter to continue straight on... but I don't know the area too well, and there are all these little strange impasses there. So, I will go to Rose Street. But when

[1]As you may remember from section 1.1. I limit my study as to what kinds of planning of action I consider, and this also shows up in the examples I am giving. I focus on a single individual's planning of *his own actions* and not on an individual's making plans for other individual's actions and also not on collaborative planning.

I get to that big play-ground, I will not walk around it but try to pass through it as a short-cut to the hotel.

This example illustrates a person *imagining* his walking back to his hotel - or 'takinga mental walk on an intemal map' - before he actually sets off. Thus, before starting to act, he does in this way decide on a *structure offuture choices,* namely, where and where not to turn.

(3) A woman has been out dancing in the town nearby and now feels very tired. She absolutely wants to catch that last bus home but realizes that there is not much time left. She might be able to do it, she thinks, if she walks through the Botanical Garden. She goes in there. When she arrives at the other side, the gate is locked though. And the walls have iron spikes on top. She thinks: If I go back now, I won't make it.. - Do I have any money fora taxi? No. Could I go and see my friend who lives in town and ask them to stay there?...no it is really too late to do that. I just *have toget* home...That bicycle back there, if I get it and lean it against the wall... could I stand on it and somehow jump over the wall?...I will try that.

The goal that this woman has is simply that of getting home. Some of the actions she proposes to herself she discards, but then she assents to the idea of trying to jump over the wall. In this sense the sequence of actions she performs when - and if - she actually goes back to get the bicyckle and so on is *prestructured.*

(4) A man comes walking through town late at night. From a distance he sees a couple of youngsters approaching. They are loud, some have sticks, and one of them points in the direction of the man. He thinks about what to do: If I walk quickly now, I may get to that building before they do and go inside, like if I were going to see a friend. If they follow me, well, then I just ring a bell. But if we meet before that, hmm - he feels in his pocket - well, I have this knife, just *in case.*

This illustrates how a man who finds himself in a threatening situation makes a plan to deal with this threat. Although there are certain elements in this plan that have a provisional or hypothetical character - like *'in case* we will meet and they get violent, I have this knife', it is clear that the man is going to take *some* steps to try to avoid what he sees as a danger.

(5) A person has just finished supper, is leaning back in the sofa and listening to some music. She thinks about the following day. If the weather is nice she will take advantage of it, perhaps go swimming?...No, rather take a bicycle tour through the woods and eat lunch there. What is there in the fridge? Eggs, sallad and some cheese... She can make a sallad tomorrow morning before leaving and decides to make some ice in the freezer before she goes to bed to keep the salad cool during the ride. She also thinks of passing the little bakery to pick up some fresh bread before heading to the woods. But it is Sunday tomorrow. The bakery opens late on Sundays..Well, it would be nice of course to set off relatively early toget more of the day. But really, she'd like some fresh bread...There are some advantages in sleeping longer as well...Yes, she decides that she will not set off very early.

This example shows a planner who is dealing with a future goal, namely, the goal of going on an enjoyable bicycle tour the following day. (She does not, we suppose, want to go cycling right now.) The person proposes certain actions that will bring about this goal and compares alternative actions as to their advantages and disadvantages. She assents to some actions and dissents to others. Most of the proposed actions will not be performed until tomorrow when she sets off on her tour. One exception is the action of putting an iceblock in the freezer (which we can call a preparatory action).

(6) A person is visiting an unfamiliar place and is going to get the keys to a friend's flat from another person. However, the visiting person feels very hungry and thinks: Is there no hamburger place around in this area? Or a shop? What if I go back to the center first just to have something to eat, but no, then I might miss the appointment to get the keys and be without a place to stay tonight. I'll go and get the keys first and just keep my eyes open for some place to eat on the way.

This example illustrates a person who tries to plan his actions in order to satisfy two goals - getting something to eat as soon as possible, and securing a place to sleep tonight. The goals conflict in the sense that he may not be able to reach them simultaneously.

(7) A person is lying on the sofa. It is a warm summer day. She thinks about the bad heating in the room. Now, during the summer, it is no problem, but sooner or later she must do something about it. At least she could buy some insulation and fix those windows. And when could she do that?...Why not take the opportunity already next week when she is going to get some material for building that closet anyway. And so she could do those windows next weekend. Or what about waiting until her daughter comes to visit who knows these things? No, it can't be that difficult; she'll try to do it herself.

This person has the goal of having hetter heating in a room of hers, at least to have it when it starts to getting colder. She thinks of how and *when* to fix this.

(8) A person is at his job. He thinks about what to do for the rest of the day since he finishes early. Yes, he wants to go to his summer house. If it is not too windy, he may take the boat and go fishing. Otherwise, he may go for a long walk in the forest and also get some wood to make a fire. It might get cold at night. What if...- the thought strikes him - when he returns, there would be someone there...Sure people break into summer houses...If he comes back from the walk in the forest and discovers this - sees that there is a light on in the house...Well, he will go into the shed and get something to...What is in there...the fishing equipment, paints, brushes, the - he will take the sledge-hammer - in case. Then the burglar comes out, stands on the staircase, he has taken the old wall clock. He'll first try to speak to him, just tell him to leave and leave the clock behind; in that case he will not tell the police. But if the burglar gets violent, he will threaten him with the sledge-hammer, to defend himself...

This final example is an example of planning that is close to imagining and daydreaming. The person thinks of the *hypothetical* situation where he meets a burglar in his summerhouse, and where the goal is to handle this situation with as little trouble as possible.

With these examples in mind, I shall now try to give a general definition of the planning of action. Among common sense conceptions of what it means to plan one's actions, one finds: 'to plan is to think beforehand about what to do and how to do it', 'to plan is to think about or decide on a course of action before acting', 'to plan is to try to structure future behaviour.' From these conceptions we can abstract the following general definition of 'planning one's actions': to have a representation of a goal and of a start situation, and togenerate one or several representations of a partially ordered set of action instructions for oneself, for getting from start to goal. The *representations* (of goals, actions, etc.) are *internat* to the planner and independent of any communication or expression of them (even if there *may* also be extemal representations *involved.)*[2]

An *action instruction* can either have the form of specifying the *action* to be performed (walk left at the square, lean the bicycle against the wall, raise your hand, put the pizza in the oven, open the door, take the keys, put the iceblock in the freezer, etc.) or of specifying a *subgoal* to reach, but where it is left unspecified as to exactly how one would do so *(get hold of* a nut-cracker, get over the wall *somehow,* walk until you come to the traffic light, make him see you, wave your handuntil he sees you, get rid of him, make something for <linner, get some bread tomorrow, etc.). Action instructions can thus be *more or less specific.*

The *generating* shall take place *before* one *possibly* attempts to act upon one's representation or plan. Planning is a temporally separate phase that is independent of the possible execution of a plan. The planning of action does not *necessarily* lead to action. Furthermore, I require that. in planning there is an *assent* or *dissent* to the represented actions or action-sequences that are generated.

This definition of planning entails that a system capable of planning its actions must be equipped with the following:

• A capacity to represent possible *situations* - in particular, *goal* and *start* situations;

• A capacity to represent and evaluate *possible events* - in particular its own possible *actions:* to represent their prerequisites and consequences, or, in other words, the situations they can transform and how. To do this, the planning system must have some representation of the *spatial and causa! s t'ucture* of its environment;

• A capacity to manipulate ideas prior to acting upon them, that is to *represent* actions and goals *without immediately acting upon them.*

The definition is intended to be a relatively *general* definition of planning. For instance, I neither require that a planner can communicate its plans or give any verbal reports at all on its planning, nor that it has to be conscious of its planning. These requirements sometimes figure in the psychological literature on planning (see, for instance, Chapman, 1984; Meacham, 1984, Oppenheimer, 1987.)

3. Routines, Programs and Plans

Planning, when sucessful, results in a plan. To rephrase the definition in the preceding section, a *plan* is a representation of a goal and a start situation and a partially ordered set of action or behaviour instructions for 'moving' from start to goal. Furthermore, this representation must be *generated* and *tested by the planner* before - possibly - acting according to the plan. When one *acts upon a plan,* the plan *regulates one's behaviour;* that

[2] The notions of representation and intemal representation were further discussed in section I.3.

is, at some level of generality it regulates *which* behaviour elements are performed and *in which order.* In other words, the behaviour elements and their order - the particular *pattern* of behaviour - is partly *due to* this *previously constructed* plan and is *not* only *due to* the behaving system's perception of the situation and the information this evokes in the system *moment-to-moment.*

However, there are many other ways of regulating behaviour that may also produce *'goal-directed'* or *'planful'* behaviour. In order to further characterize *planned actions,* we shall consider some of those forms of behaviour-regulation. Let us start with the following examples of kinds of action or behaviour that are not regulated by plans in my sense:

• *Kinetic, phobic and taxic behaviour:* These kinds of behaviour are the most primitive in the hierarchy of Lorenz (1973). For instance, there is the taxic behaviour of the flat-worm that turns until the current that brings the smell of nutrition is equally strong on both sides of its head and then starts creeping forwards until it reaches the food. (Lorenz, 1973, p.74)

• *Jnstinctual behaviour:* For instance, consider the mosquito which whenever it needs nutritional supply and perceives a particular smell and a particular temperature, performs its motor pattern for biting. Here we find the three elements of an *instinct* as a mechanism for behaviour control: a drive, a set of perceptual detectors for information that is relevant for this drive and a preformed motor pattem (cf. p.22.).

Another example is the well studied wasp *Sphex,* taking a paralyzed cricket into her burrow. Upon bringing the cricket to the burrow, she leaves it on the threshold, goes inside to check that all is well, emerges and then drags the cricket in. In the burrow she then lays her eggs which batch and the wasp grubs feed off the paralyzed cricket. During normal circumstances acting upon this routine produces adaptive behaviour. But, if the cricket is moved a few inches away, while the wasp is inside inspecting, the wasp will then just bring the cricket back to the threshold and repeat the procedure of entering the burrow to see that everything is right. This may be repeated again and again.

• *Automatized* - 'well-learned', habitual *-behaviour:* This kind of behaviour ranges from simple motor or action pattems like walking, swimming, throwing and catching a ball, breaking a twig, jumping, bouncing a ball, picking blue-berries, etc. to more complex habitual behaviour, like perfonning certain exercises every morning, following a particular path to work, greeting a king in a particular way, using a particular procedure when washing one's hair, etc.

Characteristic of all of these kinds of behaviour or actions is firstly that they are *re-active* in the sense that they are performed under, and in response to, certain comparatively specific environmental conditions. Often, they require the presence of these particular conditions in order to occur. Secondly, they are relatively *stereotype.* They are *allftxed behaviour patterns* that are repeated in just the same way under the given environmental conditions. The generic term that I have decided to use for such behaviour is *routine behaviour* or *routine actions,* and a corresponding underlying control mechanism I call *a routine.* A routine action is a fixed sequence of behaviour elements which are more or less 'automatically' set off in a situation (defined by the system's perceptual environment and motivational state). Once started, a routine action is more or less 'automatically' executed. It is a *unit* and is performed as such. Often, an observer would describe the system that is performing a routine act as 'continuing to do the same thing, performing the same action'. In executing a routine, the agent has roughly the two altematives *'continue'* or *'interrupt'* but not the possibilities of 'skipping some components', 'postponing some components' or 'reordering the sequence', etc. Toere may be some variance in the particular movements, as to speed, effort, etc., due to sensory information about the environment that time. Or, in other words, there may be what I call *localjlexibility.* But there is never any alteration of the overall pattem of the behaviour.

A routine action is characterized by *the sequence* of behaviour elements and by an *end condition* - which is either (the performance of) an *end behaviour* (like the last movement in the morning exercises) or (the reaching of) an *end situation* (like being behind the wall in an action of hiding). See the figure below.

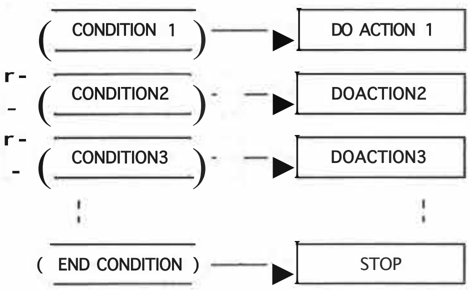


Figure 1. A routine

Of course my notions of a routine and a routine action are *encompassing notions.* Routine actions vary from *ve,y* fixed and inflexible pattems like reflexive behaviour and fixed innate motor patterns that can unfold even in total absence of any sensory feedback (Lorenz, 1973, p.79., p.83., speaks of 'Leerlaufbewegungen') to complex habitual behaviour with ample space for modification and variance due to sensory feedback, in other words, space for local flexibility. Yet the properties of boundedness to a specific environmental situation (reactivity) and stereotypicality- at some level- are characteristic of all kinds of routine behaviour.

Turning to kinds of behaviour that are less stereotype and less bound to specific circumstances, let us next consider what I call *program behaviour* or *program actions* - actions governed by programs. A *program,* in my conception, is defined by a representation of an *overall goal* Ganda *structure ofaction inst'uctions*. These can be either in the form of the specification of a *subgoal* (G' 1, G'2..) or in the form of the specification of a *particular action* (al, a2..) ora combination of these two. (See figure 2.)

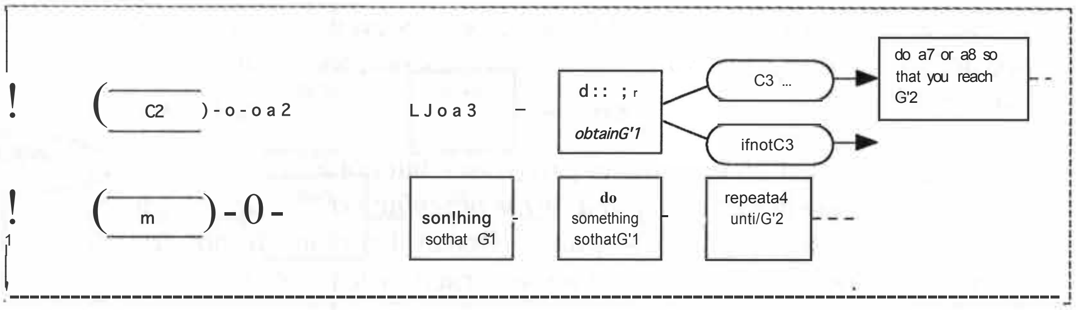


Figure 2. A program

Program is a hierarchical notion. A program can have parts that in turn are programs, i.e., *subprograms.* Programs can vary according to their generality. They can be more or less specific. (See figure 3.)

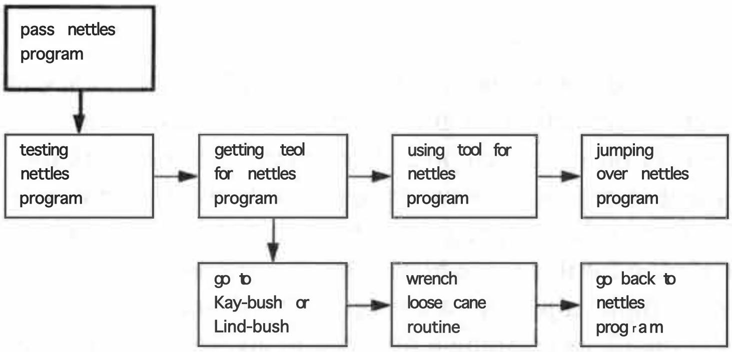


Figure 3. Hierarchical structure of programs

A program contains 'branchings' or alternatives in two ways. First when there are alternatives *given* in the program (like 'if open area, then perform a1, if limited area, then perform a2'), and secondly, when only a subgoal is specified but it is left open exactly how to reach it (like 'jump over the nettles', 'go to Kay-bush' - in some way or another). The *overall goal of a program* can be reached in more than one way, and the particular path taken can vary between program executions or actions depending on environmental conditions. A program does not specify *one* fixed *sequence* of actions for coming to the final goal, like a routine does. Furthermore, it does not have to be - and usually is not - specified all the way down as to representing particular motor patterns. The program should be conceived of as the *entire structure:* the program and a corresponding program action are defined by the represented overall goal together with all specified instructions.

It should be noted that the distinction between program behaviour and routine behaviour is not *sharp.* It is possible to *choose one* particular path through a program structure and to fill in the details down to the level of particular movements. In this way we obtain a routine. And some *routine actions* are what I call *automatized program actions.* These are actions that are stereotype and uniform, as long as the relevant environmental conditions are 'as usual' or 'as expected'. In unusual or divergent situations though, when something unexpected occurs, the acting system may utilize other alternatives than 'just continue' or 'interrupt'. It may fall back on the program instead and may thus have some chance to adjust its behaviour. (Or the actor may even without environmental variation decide to introduce some variation in his behaviour, for instance to modify his morning exercises.)

Now, *a plan* for action is a kind of *program* - but not all programs are plans. Only programs that are *generated and tested by the planning system, before* it possibly acts according to this program, qualify as *plans.* (Recall that plans do not *have* to become executed.) *Preformed* programs, be they inherited or learned (by conditioning, imitation, verbal instruction etc.), that are not generated and tested but only *activated* in the situation do not count as plans.[3]

In a program action regulated by a *preformed* program, the various alternative actions are *given* at the outset. The program structure with its various alternatives for acting is available at the moment the organism finds itself in the problem situation. In contrast, *to plan* means to *construct* such a program structure, i.e., to generate and test altematives; discard some, assent to some, etc. Consequently, *planned* actions form only a subset of program actions.

That a plan for action is a program should not be taken to mean that a plan in general can be *executed* and *by itself* causally engender action. In general, plans and planning are not sufficient to produce behaviour, but action structure comes from plans *together with* routines, habits and moment-to-moment interactions between agent and environment. It is not the case that a given action must be exclusively regulated by *either* a plan *or* by a routine or a preformed program. The reason for this is the *hierarchical* structure of behaviour and the control of behaviour. For example, a person in a shop can be described as doing shopping, as getting some milk and bread, as grasping a milk bottle, as stretching out her left hand, and so on. Therefore, even if one's main concern is with *planned* behaviour for some high-level goal, the sequence of actions can be broken up into smaller elements eventually leading to action elements that are regulated by *routines* or *preformed programs. To plan* means to organize one's behaviour - its elements and their order - by specifying and structuring behaviour at *some level of generality.* But a plan for action does not generally specify every detail down to the level of muscle contractions. For instance, in devising a plan for how to get a piece of fruit high up in a tree, by getting a stick and throwing it at the fruit so that it falls down, and then eating it, nothing concerning *how* to perform the throwing, *how* to grasp the fruit, *how* to bite the fruit, etc., has to be represented in the plan itself, because at this level the behaviour can be and is best handled by routines.[4] Or I can plan to go shopping and plan where to go and what to buy, yet not plan how to move in the different stores.

Thus the fact that an action is planned and regulated by a plan does not mean that there are no routines and preformed progams involved in the actual production of the action. (On the other hand, if an action is regulated by a *preformed program,* there can be room for

---

[3] But it is of course possible that a certain action or behaviour-pattem has at one time been planned, and that the behaviour pattern is then *learned* so that the plan is stored to be just retrieved or activated in similar situations. You may even create a routine out of a plan.

[4] It is however *possible* to plan, say, *how* to throw something, or to plan the biting and eating of an apple in detail. Think of the athlete or of the actor. - Sometimes details of automatized program actions can be planned, even though they normally are not.

planning *within* this ready-made program. One may plan how to reach a specified subgoal and one can think through given altematives. Remember that a program for action can be more or less specified or detailed).

To sum up, a *planned action* is an organization of ordered behaviour elements which are, at *some leve/ of generality,* due to *aplan* and not merely to the behaving system's routines and preformed programs and/or to its perception of the situation and the information that this evokes in the system *moment-to-moment.*

Sometimes planned behaviour is conceived of simply as behaviour directed by a *goal representation.* I shall now explain why this requirement is not sufficient to characterize planned action in my sense, and why it is necessary to require *internat generating and testing* of action instructions as well. First, just the fact that a system forms and uses *some representations* of its immediate environment and of its own actions does not imply that the system is capable of planning. *Such* an ability must, in my view, be attributed to all organisms that are capable of some adaptive behaviour. The argument for this is the following. The environment does not come to an organism 'partitioned' in one particular way. On the contrary, the number of possible ways of identifying potential 'objects', 'properties' and 'events' in an econiche is enormous. Somehow the organism has to *categorize* salient aspects of its environment. One will not get very far with psychophysical mechanisms that just transduce aspects of and events in the environment. Even relatively primitive organisms apparently recognize and respond to more complex features than just one or a few physical parameters. Information is organized and re-presented in various formats. And where this occurs, there is, I believe, reason to say that there are intemal *representations* involved.

Second, a planned action is also not sufficiently characterized by the requirement that one of these internal *representations* should be a *goal* representation and that the action is initiated by such a goal representation. *This* feature is the hallmark of a more extensive group of actions than planned actions, namely, what I call *intentional actions.* An intentional action is basically an action that is initiated by and to some extent further guided by an anticipatory representation of a certain goal. When intentionally grasping a fruit, say, or running towards a tree, it is (the activation of) an anticipatory representation of the sensory feedback response of the movement that *initiates* the movement. And this representation also *guides* the action to the extent that there is a continuous *comparison* between the goal representation and the perception of the actual position of the limbs or body parts at a particular moment in time, giving the basis for a continous refinement and regulation of motor movements. This is a kind of' *testing* one's actions'. More precisely it is a testing of the *consequences* of an action as compared to the represented goal. But the testing in *planning* is of another kind. Planning involves *internat* generating and testing of *potential* actions. It is *representations* of actions and their consequences that are generated and tested, relative to one another and in relation to an overall goal. In brief, there is a testing of actions without acting. See figure 4.

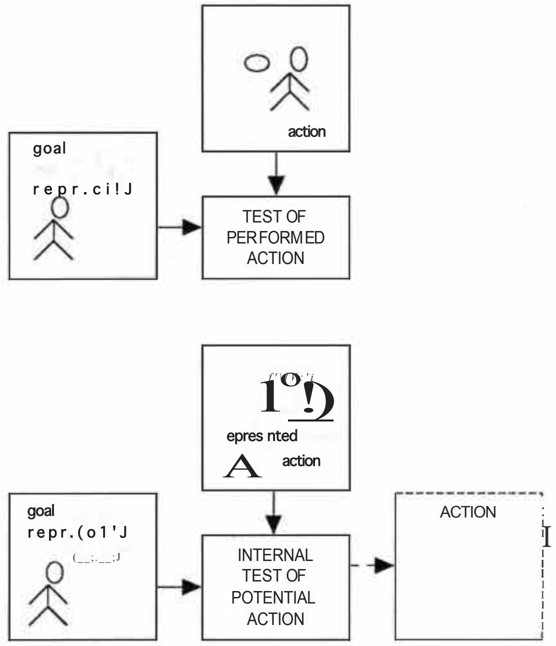


Figure 4. Testing in acting and testing in planning

To put it slightly different, in *intentional* but *non-planned* actions, all *comparisons* are made *during acting.* They involve some state actually at hand and brought about by action, and the perception of *this state* is compared to a goal representation. Whereas planning involves purely *internal* comparisons between represented and predicted conseqeunces of potential actions and represented goals, etc.

Furthermore, because planning is a question of *internalgenerating and testing,* there is an obligatory *time-lag* between the moment when a goal representation is activated and the moment of action, i.e., of the execution or the plan. It takes time to *evaluate* potential actions and subgoals and to mentally consent to behaviour elements and behaviour pattems before possibly acting. For intentional but non-planned actions there is not this hiatus between the goal representation and the motor discharge or action offset. In contrast to the case of planning, there is no evaluation of representated subgoals or actions, but rather the activation of the goal representation *automatically initiates action.* (See figure 5.)

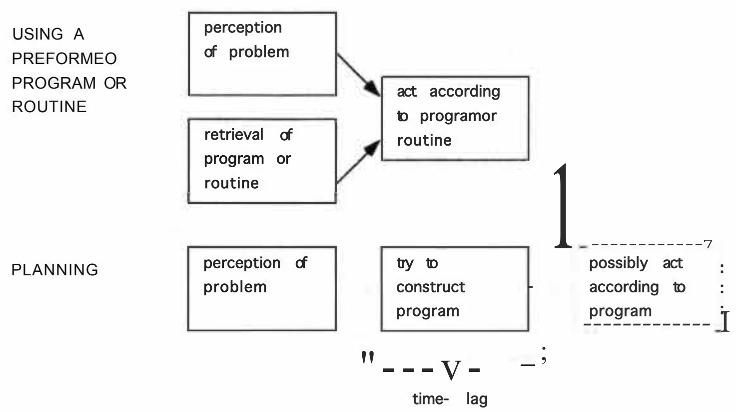


Fig 5. Planning takes time

To conclude this section, note that my conception of planning and planned actions is relatively strict compared to many definitions and conceptions of planning found in the literature. One main reason for the prevalence of broad definitions is, I believe, the influence of the definition that Miller, Galanter and Pribram gave in their seminal book, *Plans and the Structure of Behavior* (1960). They define a plan as "any hierarchical process in the organism that can control the order in which a sequence of operations is to be performed" (p.16.). Thus the origin of the program is of no importance for them. According to this definition, all program behaviour as well as much routine behaviour is regulated by a plan. However, as I have tried to make clear in this section, plans should be separated from routines and programs.

4. Recognizing Planned Behaviour: Criteria for Plan Ascription

A planned action, recall, (p.51.) is an action, the organization of which - i.e., the behaviour elements and their order - at *some level of generality* is due to a *plan,* and not merely to the behaving system's routines and preformed programs and to its perception of the situation and the information that this evokes in the system *moment-to-moment.* But how does one recognize that an action has been planned? According to my conception of planning, the capacity for the planning of action is independent of awareness about one's planning as well as of a capacity to *tell about* one's planning and plans. Thus I cannot rely merely upon reports from planning systems on their planning as criteria for planning, but instead, I choose to focus upon *behavioural* criteria. Two important behavioural criteria for telling whether an action is planned are *flexibility* and *novelty.*

That an organism behaves *flexibly* means that it can reach or maintain a certain state or situation under different circumstances by adjusting its behaviour to those circumstances. This criterion can be compared to Anderson's 'litmus test for planning' which also focuses on flexibility:

> [A] litmus test for planning [is as follows]: the system sets forth a sequence of intended actions, nates a conflict in the sequence, and reorganizes it. If it can be shown that a system reorders a preferred sequence of actions in anticipation of goal conflict, then that system is engaging in planning. (Anderson, 1983, p.167.)

A *novel* action is simply an action that has not previously been performed by the system. The 'novelty-criterion' does not of course require every element of the performance to be 'novel' but only that *at some level of generality* there is novelty in the organization of the behaviour.

In particular, the criteria of flexibility and novelty can be used to distinguish planned actions from a *routine actions* since the latter are characterized by stereotypicality, reactivity, and limited variability. Furthermore, the criterion of novelty can be used to distinguish planned actions from non-planned program actions. However, it should be noted that in order to look for and estimate degrees of novelty and flexibility of behaviour, one has to study the behaviour of a system (or a group of systems) during some time and in a variety of situations. I should also add that novelty and flexibility are no more than *signs* of planned behaviour. They are neither necessary nor sufficient aspects of actions that are planned.

Conceming novelty, a planned action does not *have to* be novel and original. This is particularly clear in cases where the planning of action involves choosing between a set of known ways for obtaining a goal rather than finding a way to obtain the goal at all. For example, assume that a person has aset of preformed programs or routines for getting home from work: she may ride her bike, take the bus or carpool with a colleague. The person is now planning how toget home on a particular day. She think through the alternatives, and finally decides to take a bus. In this example, various preformed programs for action *are internally generated and tested (evaluated).* When observing the person, however, her behaviour does not appear particularly novel. On the other hand there can be *novel* behaviour that is *notplanned.* Chance factors, trial-and-error, etc., may produce novelty or variance ofperformance.

The criterion of flexibility has more weight. Yet, preformed programs or a rich set of routines can also account for flexibility. Indeed behaviour that seems to be planned as a whole and by recognition of means-end relationships may be modelled by what is called a production system. Such a system consists of independent isolable rules which are like routines. (See, for instance Anderson, 1983; Agre and Chapman, 1978.) On the other hand, *nonjlexible,* rigid and non-adaptive behaviour *can* of course be planned. To wit, flexibility of behaviour never follows directly from the fact that an action has been planned but depends also on the *execution* of the plan. The flexibility is determined by the system's capacity for monitoring and adjusting the execution of its plans and by its capacity to decide whether the execution of a plan produces the expected results, to decide what portion of the plan needs to be executed next, to decide whether it is necessary to do some replanning, to adjust the representation of the problem, etc.You can imagine a rigid plan follower with little or no such flexibility in plan execution. Such a system may perform actions that are not flexible, in spite of the fact that they are planned

To assume that planning underlies some behaviour, i.e., to ascribe the property of being planned to an action by means of behavioural criteria, is of course a question of *interpreting* the behaviour. What is involved is an attempt to *explain* behaviour - in particular novel and flexible behaviour. Now, for any behaviour or action there are *in principle* several -indeed innumerable - models for interpreting it and explaining its grounds. At large this is independent of how novel and flexible the action is. As I mentioned, an action that seems to be planned by recognition of means-end relationships and conceived of as a whole can yet be modelled by a production system. And any behaviour may be explained in a behaviouristic framework in terms of a history of conditioning (reinforcing). Furthermore, any behaviour, however flexible and novel, *can* be ascribed to chance factors. But explanations are more or less plausible. Explanations in terms of conditioning, for instance, get increasingly strained as more novelty and flexibility become apparent in the behaviour. (For a discussion of this see Dennett, 1978, pp.67-70.) There are cases where, to my mind, none of the just mentioned kinds of explanations are plausible but where the most plausible interpretation is to assume that the behaviour in question relates to some 'prior thinking' in the sense that the agent has had a problem representation with a start and a goal, and in relation to this has generated a set of action instructions. In brief, this is to assume that a structure of choices has already been constructed. By making such an assumption we can, for instance, explain why an agent makes a *detour* that is *novel* and *appropriate.*

What is central is the *plausibility* of the explanations proposed. Regarding explanations that refer to planning, this means being concerned with whether the assumption of planning underlying the behaviour seems required and whether it produces a plausible explanation of the behaviour.

Finally, there is yet another kind of behavioural sign that I want to mention, namely, when an individual, at a decision point or just before, displays *hesitating* behaviour. The individual may stop for a while or look in several directions. It *looks as if* it is 'thinking' about where to go, which way to choose. This 'sign' may not carry much weight in isolation, but in combination with other criteria it does, to my mind, definitely add to the credibility of interpreting some behaviour as planned.

In conclusion, it is of course important to recognize the difficulties and traps involved in attempts to read out mental processes. This is however no ground for going to an extreme and regard mental phenomena as a realm completely out of reach for any scientific investigation. True it *is more controversial to say* that an animal is planning or that it has planned the action it is now carrying out than to say that an animal is now approaching a ball, or that it is located at instant t, three meters East of the berry tree. But, as Menzel (1987) suggests, the difference between the statements above is perhaps a matter of degree rather than of kind. There is no unique way to determine where observables leave off and inferences begin, and what objects or entities or processes are directly tangible and perceivable.

5. *Immediate Planning and Anticipatory Planning*

The most central distinction in this thesis is between *immediate* and *anticipatory* planning. This is roughly the distinction between planning as related only to *immediate* needs and interests and planning as related to potential, *anticipated* needs and interests *as well*. An *immediate planner* is a system capable of immediate but not of anticipatory planning, whereas an *anticipatory planner* is a system capable of anticipatory as well as immediate planning.

Two things concerning immediate and anticipatory planning should be noticed:

(1) The distinction is of interest as such because it is what I call a qualitative distinction.There is *not* just a *continuous* transition between immediate and anticipatory planning. The development of an immediate planner into an anticipatory planner requires definite design. changes. It is *not* the case that the immediate piånner possesses. all the• components required for anticipatory planning and that these just have to be elaborated and executed more flexibly or efficiently, but a real reorganization of elements and the addition of new components is required.

(2) The capability for anticipatory planning seems to be a uniquely human capability, i.e., a distinctive characteristic of humans in contrast to other biological creatures with some planning capacity. Roughly, humans alene can plan their actions not only in order to satisfy immediate needs and selve current problems but also in relation to potential, predicted needs or problems.

I define *immediate planning* as planning that relates directly and only to a *current or immediate interest and to a corresponding motivation for action.* It relates to current interests in the sense that all *goals,* including the final or the *overall goal(s), represented by the planner* relate to a current interest. There is a current problem and the planner has a proclivity or preparedness to deal with this immediately. It is motivated to *act* to solve the problem and reach the final goal *as soon as possible.* As soon as it has its plan, it will attempt to execute it. For instance, an individual is threatened and planning bow to get away from the threatening situation. Or an individual is hungry and plans how to get something to eat. (Of the examples above, pp.43-45, the first four examples are examples of immediate planning.)

In contrast, *anticipatory planning* is planning that does *not* relate *only* to a current interest and a corresponding immediate motivation for action, for instance when one is not currently freezing but thinks of the coming night that will be cold and thinks of bow to handle this, or when one is not under threat now but plans to build a trap for a potential threatening enemy, or when one is not hungry now, only very tired, but makes a plan for actions for dealing with the tiredness *as well* as the anticipated hunger. (Amongst the examples on page 43-45, the four latter examples are examples of anticipatory planning.)

In anticipatory planning, the overall goal (or one overall goal) that the planner represents relates to a potential future problem and interest and to some *actions* for which there is *no current motivation.* For instance, a plap.ner is not currently motivated to do anything about being cold - this is not a current problem - but makes a plan for actions to deal with some anticipated freezing. Or a planner is not interested presently in dealing with a threat, as there is none at hand, yet makes a plan of action 'in case that...', or a planner is currently not motivated to sleep, yet makes a plan to arrange to be able to deal with this interest in the future.

Anticipatory planning as well as immediate planning may involve the planning of actions that are to be *executed immediately,* in particular, preparatory actions such as collecting firewood beforehand or arranging for a place to sleep later on when the time comes. But in such cases of anticipatory planning, *certain other actions* - those that are *prepared for* - will not be performed until later when the motivation occurs. The *overall goal* of the plan, as concevied by the planner, is not just to have that firewood or to have those beds arranged now, but the plan relates to a future situation and some action for which there will be motivation only then; like *make a fire when* one gets *cold,* to *go to sleep when* it gets late and one gets *tired.*

Anticipatory planning can result in plans that produce or govem *anticipatory behaviour,* such as setting up traps, or obtaining and storing food. By anticipatory behaviour I mean that what is done at a particular moment in time is not biologically useful at that time in the sense that it is not required for satisfying an immediate need or interest. Yet the behaviour can *prepare* for future actions that will be biologically useful at later moments of time.[5] However, there is much anticipatory behaviour that is not due to any anticipatory

---

[5]By 'prepare for future action' I mean that something is done to establish preconditions for this action or in some way facilitate its execution. There is mental or *internat* preparation. Take for instance the skier who imagines his next run, or the person who prepares herself for a particular sight. And there is

planning. Many creatures - insects, birds, rodents, etc. - instinctively engage in food storage. And animals may migrate and set off to warmer places before it is cold, and so on. The factor that characterizes anticipatory behaviour produced by anticipatory planning is that the system really has a conception of its future interest. And thus it knows *why* it does what it does presently in a preparatory action. If the planner would have his conceptions changed due to some new information concerning an anticipated problem and interest and so for instance believe that 'alright, food will be provided during that trip' or that 'I'm not going to be tired this n ght anyway' or that 'there will be no danger after all', etc., this will influence the individual's planning and the corresponding anticipatory behaviour. (In this sense, the planner is 'free' to engage in the anticipatory behaviour or not.) In 'pure anticipatory behaviour' on the other hand, the system need not have any inkling of why it is performing these (anticipatory) actions.

This opens up some possibilities for testing whether or not there is anticipatory planning behind some anticipatory behaviour. By testing whether the individual's behaviour is sensitive to changed conditions and/or new information related to the predicted, future interest in question, one might find out whether the individual really has a conception of some overall goal that relates only to a potential motivation for action. For instance, it may be tested whether the overall goal of some collecting and storing behaviour simply is to have a supply of food now or whether there is the further goal of having something to eat in a coming situation where you will need food but food may be scarce.

Finally, given that the distinction between immediate and anticipatory planning is qualitative, a planning system *is* either an anticipatory planner or it is not. However, this does not mean that it is it always possible to definitely *demonstrate* or *determine* what kind of system a particular agent is.

6. Planning of Action and Problem Salving

I frequently use the term 'problem', saying that the immediate planner tries to solve an immediate problem, the anticipatory planner an anticipated problem. What is the relation then between the planning of action and problem solving? In much of the literature, planning is held to be something one engages in when facing a *difficult* or *urgent problem situation,* where one has to figure out bow to transform a situation, usually the present situation, into a *well-defined goal situation.* And often planning is simply *equated* with problem solving. (See, for instance, Hayes-Roth and Hayes-Roth, 1979; Miller, Galanter and Pribram, 1960; Oppenheimer, 1987.)

If we take problem solving in a wide sense, I agree that all instances of planning *are* instances of problem solving. The planner has the problem of *not knowing* exactly what to do (now, later or at a certain time) or how to do something (now, later or at a certain time) or when to do someting. *Generating an adequate plan* is to solve these problems. And the *execution of the plan,* in transforming some start situation into a goal situation, solves a problem as well. But note that *the planning of action* is only the *first* stage of a *two-stage problem solving* process, where the *second* stage entails monitoring and guiding the *execution of the plan* to a successful conclusion. (Planning is *internat* and not

---

behavioural preparation, for instance, in making a dough in order to bake some bread later on, or making one's bed in order to sleep in it during the coming night. ***Planning*** is one kind of intemal preparation. (Behavioural preparation can (in turn) - hut <loes not have to - be planned and thus mentally prepared.)

behavioural problem solving, and as such, planning alone can never finally solve a biological problem.) Moreover, the *problem solving character* of the planning varies a great deal among different instances. First, how *urgent* is it that the planner finds a plan for action, and how urgent is it that he acts? One can think of instances of planning that have the character of 'structuring or organizing one's time' or 'to have aspirations for the future' or 'to dream and play around with possibilities'. In such cases there need not be a pressing problem at hand. Besides, the goal may not be so well defined (at least not at the outset). (In general, immediate planning is more 'problematic' than anticipatory planning in the respect of urgency.) Second, how *complicated or unusual* is the situation that the planning relates to? How different is it from previously encountered planning situations? To what extent can the planner rely on ready, preformed plan-bits?

To summarize the discussion of the relation between the planning of action and problem solving, one can say that in a wide sense planning is problem solving but it can have a more or less strong character of problem solving.[6] And there is of course problem solving that is not the planning of action. Problem solving does not have to relate to coming actions or to actions at all, and problem solving may be behavioural as well as intemal.

7. *Planners in Nature - Biological Planning Systems*

In this section Ishall address the following questions: Which living creatures are planners? Which biological systems generate action instructions for themselves in the sense I have described? That is, which biological systems generate representations of sets of actions and subgoals for moving from a start toa goal situation before they possibly act according to this representation?

I take it as indisputable that *human beings are planners* in the sense that they are capable of and actually do engage in planning their actions.[7] They make plans, for instance, as how to get to town this evening, what to do until the next lecture begins, how to get home now that it has started to rain, when to take a holiday, what to do about the stain on the new cardigan, where to go to buy a hammer, how toget to the bus-station in time, how to raise the temperature in one's flat, what to cook for <linner, how toget in touch with a certain person whose number and address one has lost, etc. Toere is evidence for this not only in their behaviour (some of) which is flexible, novel and innovative and for which the most plausible explanation is that there is planning involved, but also in verbally (or otherwise symbolically) *communicated* plans; maps, schedules, shopping lists, strategies that are written down, and by the fact that people relate their plans and planning to one another.

With other creatures though, one is confined to behavioural criteria, as the ones I have spoken of in section 5.4. The evidence according to such criteria gives, I purport, enough

---

6And note that the planning of action is an appropriate strategy only when the situation is 'just sufficiently difficult or problematic'. In a wholly familiar situation there is hardly any need for planning one's actions since there will generally be preformed programs or routines for action. And in a completely novel and unfamiliar situation that the planner has difficulties understanding at all, successful planning of action is not possible.

7However, which actions are planned, bow *much* humans are engaged in planning (to what *extent* they are planners), and questions of individual variation, and so on, are separate issues.

ground to assert that human beings are not the only planners around. We *at least* share the capacity for planning our actions with *other primates* - chimpanzees, gorillas, etc. To support this let us look at some examples from various observations and studies:

(1) Fig-gathering: A group of chimpanzees is at a fig tree, where the trunk of the tree is too large for the chimpanzees to scale. But one limb of a tree nearby permits possible access. Two individuals altemately stand on this limband hit at the fig tree's nearest branch with sticks which they have broken from this tree and stripped off the bark and thoms. Later, another individual bounces the limb up and down and rises bipedally when the limb is at its peak. He can however still not reach the fig branch. After a while he begins to break off most of the branches from the limb on which he sits, now not stripping them as before but simply dropping them to the ground. The limb which now is lighter starts to rise. The individual starts to bounce the limb again, rises bipedally thus getting hold of a fig branch and climbing into the tree. (Menzel, 1987, p.60.)

This is a typical instance of problem solving behaviour that seems to be preceded by *internat problem solving.* It seems not to be *by chance* that these *two* actions and in this order - breaking off the branches and bouncing on the limb - are performed. When one ohserves this behaviour, one is indeed inclined to say that there has been some prior thinking involved, some testing of potential actions.

(2) Nut-gathering: A group of chimpanzees uses hard rocks and heavy sticks to crack and eat the hard Panda nuts, and softer clubs for Coula nuts. Although they cannot see from one tree or rock to the next, they will directly get the appropriate material, and they will go to rocks that are close to the nut tree they are out to exploit. *Least-distance* is a favored strategy, although both weight of the stone and distance seem to be taken into account simultaneously. *(Ibid.,* p.62.)

Again, the action patterns of these creatures - where they go and in what order and what tools they get - seem to be *prestructured.* If they used some more random or opportunistic strategies, the fact of their 'least-distance displacements' would be difficult to explain. (And explanations that refer to use of sensory cues or to stimulus-response-associations seem farfetched.)

(3) Termite-fishing: A chimpanzee looks for a while at the various twigs of a tree, then 'chooses' one, breaks it from the tree, strips off the leaves, and takes it back to the termite heap. The chimp then finds a tunnel, scrapes away the thin layer of soil that covers the entrance, pokes the twig into the hole, extracts it with its attached termites and eats them. (von Glasersfeld, 1976, p.217; Gardner, 1983, pp.216-217.)

Illustration 1. Thinking

Again this behaviour, as an example of novel or innovative behaviour, may strike one as 'contemplated'. Following von Glasersfeld *(ibid)* the remarkable thing is that there seems to be a novel reference item involved in this chimps behaviour, namely, an idea or representation of a tool. This is what it is acting upon: it does not tear leaves of the branch in order to eat them, as usual, but, precisely, in order to transform the twig into a stick-like tool. And it is of course not by accident that the chimpanzee, after having fabricated the tool, *goes back to the termite heap* - instead of doing something else. That 'choice' was already taken beforehand.

(4) Deceiving competitors: A band of vervets is losing ground in a territorial fight with another band. One of the losing-side monkeys then suddenly issues a leopard alarm (in the absence of any leopards), leading *all* the vervets to take up the cry and head for the trees - creating a truce and regaining the ground that this band had been losing. (Dennett, 1983, p.347.)

Did this individual vervet *think* and *predict* that such an alarm would have this effect? Is it truly *decieving* the enemy monkeys? What does it believe - if anything? There are indeed several possible explanations that ascribe to this individual vervet monkey more or less sohpisticated cognitive competences. For suggestions and a discussion, see Dennetts article (1983) where he also discusses possibilities for experimental testing of the various suggested explanations. Without such tests, however, it is the plausibility of the various explanations that is the leading criterion.

(5) Getting the bananas for oneself: There is a chimp who is part of a large group and has not managed to get more than a couple of bananas for himself. He gets up and walks away, the other chimps follow him. Ten minutes later he returns by himself to eat bananas. He performs this 'leading the group away' several times. (Whiten and Byrne, 1988, p.238.)

(6) Taking short-cuts, etc.: Here we have, for instance, Menzels well-documentated study where a chimpanzee is carried around a field and allowed to watch as an experimenter hides up to eighteen pieces of food in natural cover. When let loose, the animal finds virtually all of the food, taking a route that bears no detectable relationship to the route which the experimenter took and which is not much less efficient in terms of the

overall travel distance than that which it might have followed if all of the food had been visible at the time of response. (Menzel, 1973, pp.943-945.)

The animal obviously does not repeat a displacement that it has aleady experienced, or act according to some associative chain of perceptions. It is more plausible that it has *reorganized* its experiences or perceptions - that it has operated on some kind of *internat map.*

(7) La.dder-invention: A chimpanzee gets up in a tree passing electric wires that are wrapped around it in the following way: It denudes a long tree branch, secures its base on the ground, sets it against the top of the tree and uses it as a ladder. (Menzel, 1987, pp.58-59.)

This is another piece of novel or innovative behaviour that is hard to ascribe to chance factors and pure trial-and-error behaviour. In that case, one would not expect such direct performance of a successful solution but also some trying with less appropriate and successful behaviour. That is, one would also expect some *errors.*

In these and other instances of behaviour, there is, in my opinion, enough evidence of flexibility and novelty to conclude that these creatures have a capacity for the planning of action. I am not suggesting of course that every element of the performances is 'novel' (cf. p.53). For instance, in the example of fig-gathering, the *elements* branch-breaking, limb-bouncing, hitting-at, are all routine actions, common in the behavioural repertoire of primates. Yet there is, just as in several of the other examples, at *some level* a 'genuine' novelty in the organization of the behaviour. This is difficult to account for by invoking conditioning or chance, etc.; in particular when we take all instances of such behaviour collectively.

The further down one climbs on the evolutionary ladder, the less evident is it that there are planners in my sense. However, in all of the following species there is *some* evidence of planning: dogs, horses, cats, rats and hamsters. (See *Cognitive Processes and Spatial Orientation in Animal and Man,* 1987, chapters 5, 7, 8 and 9.)

Almost all of this evidence - as well as and much of that concerning planning in primates - is related to *spatial tasks* and to *locomotion.* Here, use of *short-cuts,* making *detours,* use of *novel paths* etc., are observed, particularly during experimental studies. Animals find original and 'smart' solutions in modified familiar tasks. They may for instance immediately take the most direct of several possible ways to reach a goal. They succeed well in detour tasks .(that is, where one needs to circumvent an obstacle or move away from the goal in order to reach it) when the goal is not visible from the choice points. And many animals can use short-cuts, again although the goal or local cues are not visible. In my view, the most plausible explanation of their occurrence is that they are due toa prior intemal generating and testing of potential actions. The individual has some problem and goal representation and in relation to this, it generates aset of action instructions for itself. And finally, also signs of *hesitation* are observed; where individuals just before reaching a choice point display hesitating behaviours, stop for a while, look in several directions, etc (cf. p.55).

What about anticipatory planning then? *Humans* can and do engage in anticipatory planning.[8] Again there is behavioural as well as symbolically communicated evidence for this. But there is no convincing evidence that any other species does in fact engage in anticipatory planning.

The squirrel is often mentioned as a common objection. 'Isn't the squirrel that gathers nuts and stores them for the coming winter an anticipatory planner?' No. This behaviour in the squirrel - like in many rodents and some birds - is *anticipatory behaviour* but without underlying anticipatory planning (cf. pp.56-57). It is routine behaviour - instinctual behaviour- that appears stereotypically in all individuals without sensitivity to varying circumstances. (In particular, without sensitivity to variation that relates to the need or interest in question, like, say, that they are continually supplied with food, that one fills their stores, etc.) There is no reason to assume that these creatures have a representation of a future interest and motivation. There is no reason to assume the representation of a potential problem for which the individual generates and tests action instructions.

The occurrence of anticipatory planning of action in creatures other than humans ought, however, in the first place to be expected in higher non-human primates. But even here, there is no evidence for anticipatory planning. None of the examples given above, pp.59-61, relates to a potential interest and problem but only to immediate needs and problems. My hypothesis is that the capability for anticipatory planning *characterizes human beings* and distinguishes them from other biological planners.

8. *The Illustrative Mode! - The Berry-Creatures*

To conclude this chapter I will introduce *the Berry Creatures.* These are fictitious creatures that will serve as an illustration throughout a main part of the thesis.The Berry-Creatures come in two versions, the 1-Creatures, and then the A-Creatures. The primary purpose of this fictive case is to illustrate the distinction between immediate and anticipatory planning. The 1-Creatures are immediate planners. The A-Creatures, evolved from the 1-Creatures, however, are also capable of anticipatory planning.

The rationale for speaking in terms of fictitious planners, recall, (pp.7-8), is that it enables me to give examples in some detail and suggest plausible properties of planners and planning capabilities without having to assert that what I say is in all details true of biological planners, *or* that it has to be true in principle of *every possible* immediate or anticipatory planner. Speaking in terms of fictive yet biologically realistic creatures enables me to discuss various possibilities that yet are constrained. Furthermore, reality is complex. This example, on the contrary, is manageable and even implementable.

---

[8] Again, to what extent and under what circumstances humans engage in anticipatory planning and whether there are individual differences are separate issues.

Illustration 2. Berry-Creatures

In 'constructing' the Berry-Creatures I have been led by considerations concerning living biological creatures and by data from ethology, comparative psychology and anthropology. In the first place, this data concerns higher primates, primarily chimpanzees, and secondly, I use some anthropological data and hypotheses concerning *early humans.*[9] The intention and hope is that the Berry-Creatures contrast with *non-fictive* biological planners, *not* by being *completely different* from those but by being much simpler. There is one *major* simplification, however, that really makes the Berry-Creatures different, at least from the creatures they are modelled after. The Berry-Creatures are *non-social.* I will return to this. (Also, when some particular details are fictional, there may exist a correspondence to the non-fictional case. An example is the following. Both I-Creatures and A-Creatures seek shelters, but I-Creatures only do so when they are cold, whereas A-Creatures may do so without being cold. The parallel to this is that early humans were presumably the first creatures to collect wood and make fire *before getting cold.)*

Their biological interests, their basic constitution and the circumstances under which the I-Creatures and the A-Creatures live are the same, and this is what I will present here. In coming sections then I will give examples of and discuss their partly differing behavioural and cognitive capacities - in particular their planning activities and capacities. Let me start with the interest hierarchy of a Berry-Creature. (See figure 6.)

[9] Tuis relates to the hypothesis that the capability for *anticipatory planning* first emerged somewhere in the transitions from apes to humans.

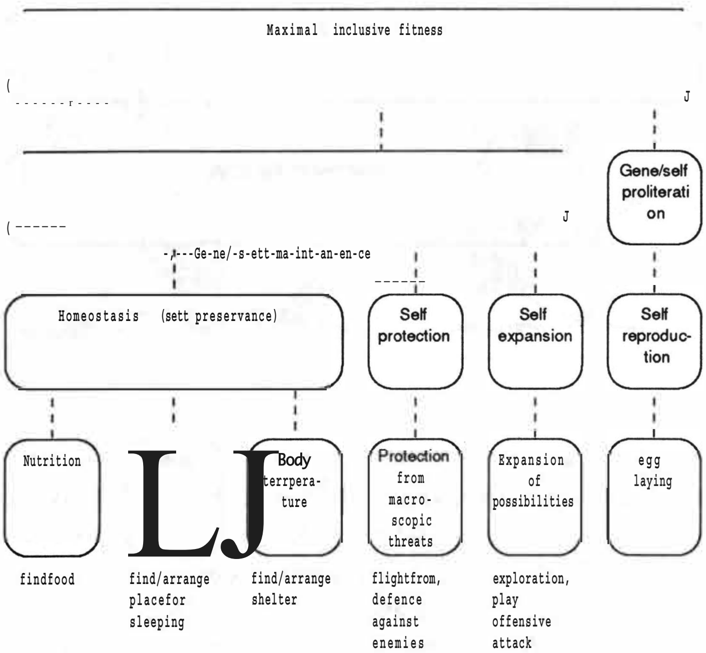


Figure 6. Interest hierachy of the Berry-Creatures

The basis is the mammalian interest hierarchy, p.24. For the purpose of simplification, however, I let the entire right part of the interest hierarchy ÷ the interests of proliferation - be straightforward and simple to deal with for a Berry-Creature compared to the interests on the left side. *Reproduction* is in principle self-regulating. There is no problem of *mating* (finding a partner, competing for this, etc.) as they do not reproduce sexually. And there is no problem of *caringfor offspring.* They reproduce by egg-laying. All eggs are hatched, in principle no matter where and when they are laid. And newly hatched Berry-Creatures can in principle take care of themselves. Toere *is* more elimination of younger Berry-Creatures than of older. The younger are *less ski/led* in climbing, in producing tools, in dealing with enemies, etc., but they are defenitely not helpless and are bom with a sufficient repertoire of behavioural programs and routines. The challenge is to *keep fit and survive.* But as long as a Berry-Creature survives and is in a certain condition as to bodily maintenance (that it is not starving or has had too little sleep), it will reproduce. Differential *reproduction* has to do with differential *survival* and not with differences in mating or dealing with offspring. Put otherwise, differences in reproduction are not due to more or less success in mating and taking care of offspring but to more or less success in surviving and dealing with homeostasis, self-protection and seif expansion - which are the interests to the left in the hierarchy. It follows from this that there is no *social interaction* within *families.* There are also no interactions within or between other kinds of groups. There is no competition between Berry-Creatures, and they do not cooperate. In short, the Berry-Creatures are decidedly *non-social.*

This is indeed is a major simplification, and I do not pretend that this is modelled after some living creatures. Remember, however, that I am focusing on one individual's planning its own actions. And so I believe the construction of the Berry-Creatures as non-social is legitimate. Furthermore, it is in principle possible to extend the illustrative case and take social aspects of planning into consideration.

Turn now to the interests to the left that are not so straightforwardly taken care of: nutrition, sleep, body temperature, self-protection and self expansion. The interests of the Berry-Creatures are matched by their motivational structure. Toere are the *primary motivations* that motivate engaging in food searching and eating, getting sleep, getting shelter and dealing with enemies. And there are the *secondary motivations,* boredom, curiosity and anxiety that work together in the way presented in section 3.4, p.27, and motivate exploration and play as well as withdrawal, rest and search for calm. Primary and secondary, remember, means that the secondary interests are not dealt with until the primary interests are satisfied. An individual will not engage in exploration or play or in just resting if it is hungry or cold or under threat. Yet secondary motivations may influence behaviour driven by primary and thus dominating motivations, in particular when these are not too strong. Curiosity rather than anxiety may lead a hungry individual to engage in more riskful behaviour to get food or make it try to eat something which is unfamiliar but looks edible.

Let me now say more about the details of these motivations and engagements in a Berry-Creature. I will in other words discuss what problems and opportunities a Berry-Creature's environment presents in relation to these interests. The world of the Berry-Creatures has the character of a forest with trees and bushes. In particular there are many dense and thomy bushes - *Nettle-bushes* - that make up hedges. These bushes grow as well as wither relatively quickly. Toere are also *paths.* These change somewhat as Nettle-bushes grow and wither away. In addition, there are *ditches* that are sometimes filled with water. On some kinds of trees there are *berries* - and this, not surprisingly, is what Berry-Creatures eat.

First, there are *Red-Berries.* A Berry-Creature gathers Red-Berries by takinga twig from a bush and climbing up the tree until it gets to the branches with the berries. These branches are thin, and too thin even for a little Berry-Creature to get out on, and as the berries grow a bit out on the branches, it is not possible fora Berry-Creature to reach thein directly. Instead it will aim at a bunch of berries with the twig that it has taken along and hit them so that the berries fall down onto the ground. Then it climbs down to get them. This is habitual behaviour in a Berry-Creature, for which a young Berry-Creature already has a preformed program. Second, there are *Shell-Berries.* These have a hetter nutritional value. The procedure for obtaining Shell-Berries is the following: Jump up into the tree, which is not so easy if the lowest branches are high up, then climb further up and use a similar procedure as with Red-Berries toget the berries down, but here with a bigger and harder kind of twig. To open the shells, the Berry-Creature then uses two stones, a big one, on which it puts the berries, anda small one, which is used to break open the Shell-Berry. Apart from the difficulty of getting into the tree, it takes energy to transport the stones to the site. They are heavy to carry fora Berry-Creature. (But they are not so for bigger creatures, and when such a creature passes, it may well kick a stone into a Nettle-hedge or down into a ditch.) Third, there are *Plumb-Berries,* round, large berries that grow very high and at wich Berry-Creatures throw stones to have them fall down. Finally there are *Good-Berries.* These berries have a very high nutritional value. But they grow on the 'outskirts' of the Berry-Creature's natural environment, and a Berry-Creature has to pass 'dangerous areas' to get there, areas where there are no trees to climb up into if they encounter an enemy. Furthermore, it is a long way to getto a shelter if it gets cold.

Toere are other kinds of trees and bushes; *High-trees, Kay-bushes, Long-Twig-bushes, Lind-bushes, Thorn-trees* and *Bast-bushes.* Bast-bushes have very long and thin, thread-like twigs, that sprout every time it rains (which does not happen so often). The Berry-Creatures use trees as a refuge to escape from enemies and some trees and bushes for getting tools. Another feature of the environment are *caves* in certain areas. Warmth that comes up from warm sources produces those caves, and the Berry-Creatures use them as *shelters* to protect themselves from the cold. In particular, when it is windy, a Berry-Creature needs to keep its body temperature from sinking - if it gets too low it will not survive. And when a Berry-Creature is cold, its habitual behaviour is to go to such an area with caves. It will usually dig a bit further to increase the size of the cave, using a twig from a Kay-bush. The Berry-Creature then gets down in the cave and covers the floor with some branches. It stays there until it is not cold any more and it is less windy, *or* until hunger or sleepiness takes over.

At night - when it is dark - a Berry-Creature will sleep (it gets tired when it gets dark). When it does so, two things can be critical. The first concems the ambient temperature and maintenance of *body temperature.* It is often windy at night. The other critical factor concems *enemies.* The *Night-enemies* are noctumal and prey upon Berry-Creatures. They are also much bigger than Berry-Creatures. When encountering a Night-enemy, a Berry-Creature has no chance to escape. To be in a cave is good for the purpose of maintaining body temperature, but in there a Berry-Creature is easy prey for the Night-enemies. However, there is a particular kind of tree - *Umbrella-trees* - that are suitable as places for sleeping. In an Umbrella-tree a Berry-Creature cannot be reached by a Night-enemy, and it is also protected from the cold. At night these trees let down their dense branches as an umbrella, and this provides protection from the wind. The procedure for entering an Umbrella-tree is the following: go to a Long-Twig-bush and wrench loose a long twig from it, take this to the Umbrella-tree and use it to bend down some branches to be able to climb the tree, and then climb up. Again, this is a routine that young Berry-Creatures already have. If it is *very* windy, however, which is the case maybe one night out of twenty (irregularly), the Berry-Creatures have problems, because then the branches of the Umbrella-trees are effected by the wind and end up in a horizontal position. The Berry-Creatures will get cold and wake up. They ususally go to a cave, a shelterplace - which of course implies a risks of getting killed and eaten. Also, if a Berry-Creature wakes up because of hunger when in the Umbrella-tree, it will leave the tree.

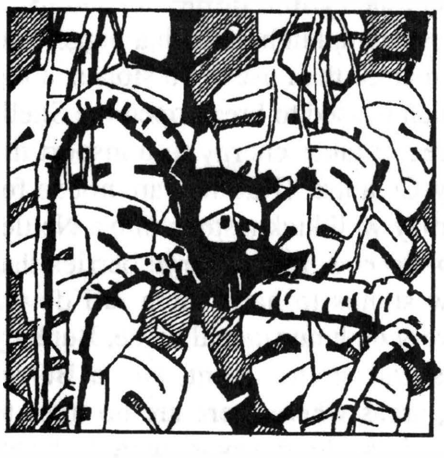


Illustration 3. In an Umbrella-tree

The main threats to a Berry-Creature are the *cold* and *enemies.* I already mentioned the Night-enemies. There are also *Day-enemies* which prey upon Berry-Creatures as well. A Berry-Creature has some chance however to deal with these enemies. They have a particular smell that the Berry-Creatures detect. The best thing a Berry-Creature can do is to escape up in a tree, as the Day-enemies do not climb. When encountering a Day-enemy, there is also a small chance of defense if the Berry-Creature can hit it hard in the stomach with a twig, so that it falls, and then escape towards a tree. A Berry-Creature can run quicker than a Day-enemy, but not for long. Day-enemies are much more perseverant.

Berry-Creatures *move around* a great deal. Food sites, shelter places and sites for sleeping are not close. When a Berry-Creature moves, it encounters various kinds of obstacles. Getting to a water filled ditch means that it cannot pass there but must take a detour. It is too broad to jump over, and Berry-Creatures cannot swim. Sometimes trees and bushes, in particular Red-Berry-trees and Kay-bushes, fall down and lie as obstacles over the path. These however are relatively easy to deal with. Berry-Creatures have programs, ways for climbing over them. The Nettle-hedges are more difficult. Yet there are ways of getting around them, ways that depend upon bow thick they are and bow broad the area is that is covered. If the Nettle-branches are not thick, they can easily be bent down or even cut off. And so, if the hedge is narrow, a Berry-Creature may just hit the branches to lower them and then jump over. If the hedge is broad, it is necessary to make a path. In this case the Berry-Creature may cut off some branches, sweep them away with some other twig, and then pass. If the branches are thick, they are difficult to cut, but might be bent down, and so if the covered area is narrow, the Berry-Creature can do just the same as with thin or young Nettle-bushes. If the area is broad, however, there is notbing that can be done. The various obstacles are obstacles for the enemies as well.

The environmental changes relevant for the Berry-Creature's primary interests are the following. There are short term changes; some regular, like dark-and-light (that decides sleep rythm), some irregular like the wind, that can blow more or less, and from different directions. When the wind is strong, Berry-Creatures have considerable difficulties walking in headwind. It is very consuming, and they get cold. Rain is also irregular. It rains relatively seldom, and then the ditches get filled with water but may dry up when it does not rain for some time. Some bushes and trees grow particularly fast when it rains. Specifically, the Bast-bushes, we know, get lots of sprouts when it rains. And there are more long term changes, what one could call seasonal changes, periods during which it, for instance, rains more often (although still relatively infrequent) than others.

By now I have related the most important features conceming the *primary* or serious *interests* of the Berry-Creatures. Apart from this, Berry-Creatures engage in play and exploration on the one hand and resting on the other. They walk around in different areas, look around, sniff at things, lift things, manipulate objects, lift stones, throw stones, climb trees, jump, etc. The benefits of this is increased knowledge and competence - which outbalances the <langers involved in these activities, like falling into a ditch for instance. And they also at times seek tranquility and rest. A Berry-Creature may rest in a cave or a tree.

To sum up and conclude this section, the interests, motivations and engagements of a Berry-Creature are presented in figure 7.

interests	*motivations*	*engagements*
nutrition	hunger	search for food, eating
sleep	tiredness	search or arrange place for sleeping
maintenance of body temperature	being cold	search or arrange shelter
protection from enemies	fear, anger	flight, hiding, attacking
novelty	boredom, curiosity	exploration, play
familiarity, security	anxiety; desire for challenge	withdrawal, searching caim and 'the well-known'; exploration

Figure 7. Interests, motivations and engagements of a Berry-Creature

6. A PERCEPTUAL-BEHAVIOURAL SYSTEM

1. Introduction

A biological system capable of planning (some of) its own actions must primarily be capable of *behaving* and of *perceiving*. Therefore my design of a planning system takes its departure from a design of *'a perceptual-behavioural system'*. It should be noted that I use the term 'design' in a fairly loose sense. I mean by 'design' a description of the underlying organization of a system that indicates the sub-systems and their properties - sub-capacities, sub-functions - and how they work together to produce the properties of the whole system.

This chapter is devoted to a discussion of such a perceptual-behavioural system. The rationale for doing this is that I consider planning capacities in a biological or evolutionary context where the planning of action is an *advanced* form of behaviour-regulation, and as such does not turn up from nowhere. It is not a capacity that just arises with its own new sub-capacities, mechanisms and substrates. But it builds upon, is integrated with and extends already existing modes of behaviour production.

My discussion of *perception* will be quite extensive because I am convinced that this is the most fundamental cognitive activity from which all other cognitive activities - planning included - must have evolved. Many of the aspects and properties that I describe in this chapter are properties that are also useful in connection with planning, namely, the *hierarchical* character of perception (as well as of behaviour), the presence of higher-order *abstract* perceptual *representations* and the *subjective and constructivistic* aspects of perception. In the two last sections, I speak of behaviour as the control of perception and of perceptual-behavioural systems of different orders. I introduce the concept of trial-and-error-behaviour that is going to be useful later on. By the end of the chapter I will have related a design of a basic system that in subsequent chapters will also be endowed with a planning capacity.

In this chapter I follow to a large extent W. T. Powers in his book *Behavior as the Control of Perception* (1973).[10] The reasons for choosing Powers' theory of perception and behaviour as a basis for my own theory of planning are the following. In contrast to many theories, it deals with behaviour as well as perception on *different* levels of *complexity*. Furthermore, it presents a perceiving-and-behaving system *not* endowed with high-level cognitive capacities but that can be extended in a reasonable way to include such capacities. Finally, it is a theory that, in spite of its comprehensive nature is thoroughfully worked through and has empirical underpinnings.[11]

---

[10] All following references to Powers are references to this book.

[11] For a presentation of theories similar to Power's, see Bruce and Green (1990).

2. *Perception and Behaviour*

By *perception* in a system, I mean the following: all processes which, on the basis of *transducers* - i.e., some channels that react to physical changes in the environment just outside the system - together with other information structures in the system inform the system about *what is* (and where) and *what happens* (and where) around it and inside it. Thus, I include in perception the entire set of events all the way from sensory receptors to the highest centers in the cerebral cortex. Note that I do *not* require that a perceptual process or a perception (a perceptual signal, a percept, etc.) is *conscious,* nor that it is *intentional* in any of the following senses: object-directed, with content, interpreted, voluntary or on purpose.

Perception is a *continuous activity.* Biological systems perceive *all the time.* Furthermore, they *behave* - they move and they do things - most of the time. *Behaviour affects perception* in the sense that by behaving the organism affects what happens around itself and within in itself, and it affects how it is informed of this.

To speak of a system's *behaving* or *acting,* I require that visible overt movements, which on the lowest level relate to muscle contractions, are involved. Planning, thinking, perceiving, feeling etc. are not actions but *activities,* which I use as a more encompassing term. Note that I use 'action' *interchangeably* with 'behaviour'. Action often is more restrictively defined, so that action is behaviour that is conscious, intended, planned, voluntary, etc. Along with this frequently goes the conception or assumption that there is a dichotomy in nature roughly of the following kind. On the one hand there is the category *'only behaviour',* namely, the movements of creatures that are *caused* much in the same way as physical events and processes. On the other hand there is *action.* Actions are not caused in this way but willed, intended, etc., and in some sense belong to another realm than 'the physical world' or 'the realm of natura! sciences'. In addition, it is not so seldom that this is coupled to the view that only humans are capable of acting, whereas other creatures behave. However, I do not build any such assumptions into the vocabulary. And when I speak of a system's or organism's behaviour or action, I'm not by using one or the other term saying anything about its intentions, planning, consciousness, etc.

Following J. von Uexkiill, an organism lives and behaves in its particular *Umwelt* (see p.22). To refer to the totality of a system's access to the world, that is, to all its epistemological resources for *perceiving* and *thinking about* its environment, I will use the term *internat environment.* This also includes all stored representations - 'memory-representations' -which do not relate directly to *current perception.* Furthermore, an organism has, at any moment of time, a *perceptual environment.* This consists of that part of its environment that is *currently perceived,* and as the organism moves around, its perceptual environment is transformed.

An extremely important consideration is that *behaviour* and *perception* are *hierarchically organized* phenomena. Regarding behaviour, an organism can, during a short period of time, be <loing, for example, all of the following: getting something to eat, gathering berries, hitting a branch on a tree with a stick in order to get the branch to swing so that ber-

ries will fall down, grasping a stick, moving its arm and hand, taking in certain body positions, and using certain museles in particular ways (see figure 8).[12]

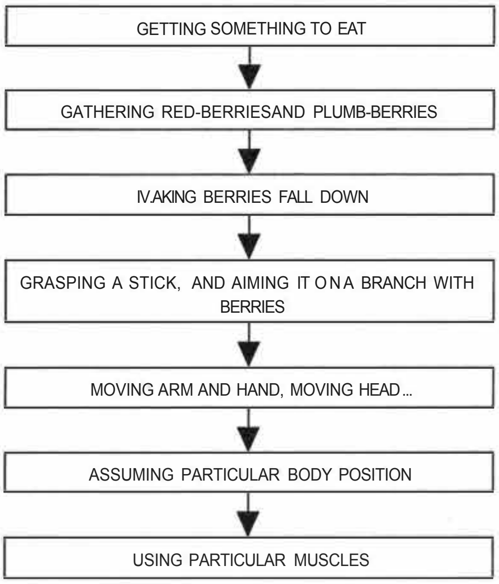


Figure 8. Hierarchical organization of behaviour

Perception is correspondingly hierarchically structured (see figure 9). An organism can, for example, at one time perceive all of the following: intensities (of light, of pressure, etc.), brownness, gloss, roundness, smoothness, surfaces, pressure (from the supporting ground), hardness, a hard, smooth surface, a hard brown berry, a branch, a tree, an arm, a moving arm, a branch moving in the wind, the grasping of a berry, biting, biting a berry, the holding of a stick and hitting a branch with it to have berries fall down from the branch, the gathering of berries, getting something to eat.

[12] You may want to object, that some of the descriptions, from a particular level in figure 8, are only ***interpretations*** of what actually takes place or what the organism is 'really ***doing'.*** True, I answer, they are interpretations but so are then all of the descriptions. And this does not exclude - quite to the contrary - that the behaviour is 'really taking place'. There is no principled boarderline to be found in the list above as to whether something is 'really done' or just interpreted as being done.

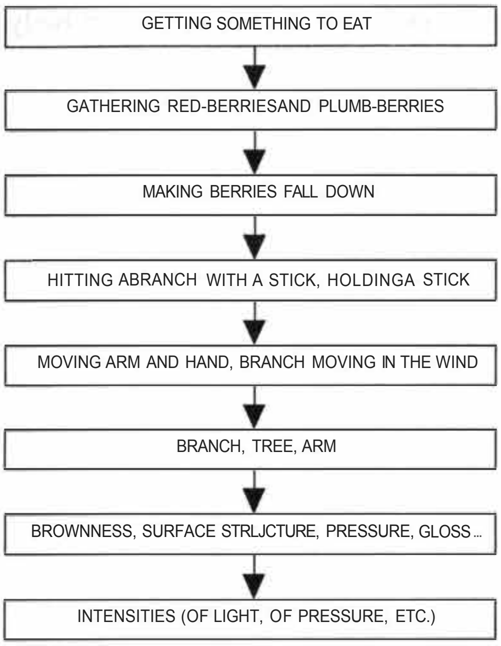


Figure 9. Hierarchical orgaization of perception

3. Perception as Hierarchically Organized

In this section the hierarchical organization of perception will be investigated in greater detail. In my model, which is a modification of Power's model, I make use of the six leveis presented in figure 10.

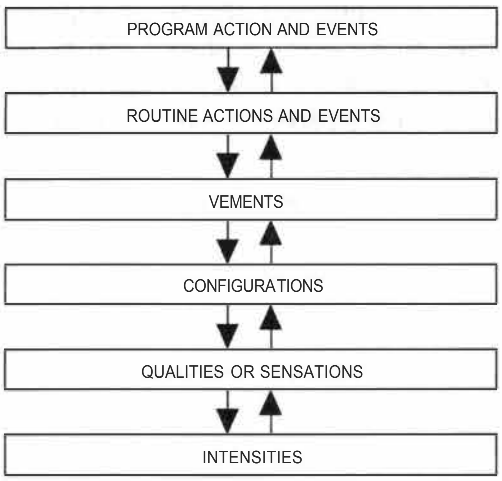


Figure 10. Perceptual orders /Levels of perception

On each level one finds *perceptual inputfunctions*. These are functions, implemented in neural structures, that can be activated to deal with particular kinds of information or input from lower perceptual levels. Signals or perceptions from one level enter certain input functions on the next level, which generate *higher-order* perceptions. A second-order perception, for instance, is the result of a perceptual input function - like a sum, a derivate, a weighted sum, etc. - of a number of first-order perceptions. On each level a great number of perceptual systems are simultaneously activated and operate more or less independently of one another. All of them receive input from lower-order systems, and partly from the same lower-order systems. Furthermore, input from *several* lower-order systems reaches one and the same higher-order system (ora set of higher-order systems). Thus, perceptions of taste, smell, temperature and efforts of biting, can, for example, contribute to the perception of an apple, or perceptions of sound and effort to a perception of screaming.

On the lowest level one finds *perceptions of intensities -first-order perceptions.* On the second level we have *quality-perceptions,* or *perceptions of sensations* - like perceptions of pressure, temperature, colour, taste, effort of biting, smell, lines, edges, points. The third level is the level of *configuration perceptions.* An organism perceives configurations of sensations; of edges, points, surfaces, colours, smells, etc. Important configurations are forms, positions, distances and surfaces. Perceptions of *objects*[13] also are the output of sophisticated third-order perceptual functions. In particular, an organism perceives its own body configurations. On level four we find perceptions of *configuration transitions* or *motion,* of rising, falling, approaching, etc. A system can perceive its own motions, as well as motion of other objects and configurations. Up to level four I follow Powers' presentation quite closely, not because I am convinced that his analysis is complete and correct but because it is not essential for my purpose to dwell on the details, in particular the details conceming lower-order perception.[14]

On the fifth level we find *perceptions of routine actions and events.* In the preceding chapter I defined a *routine action* as a sequential behaviour pattem of invariant order tied toa relatively specific situation or input. It isa stereotype behaviour pattem that is performedasa unit. One criterion of a routine action lies (cf. p.48) in the sense of 'the same thing continuing to happen'. It is stereotype in that there is only 'local flexibility', that is, certain variability in speed, detailed movements and muscle contractions, according to varying environmental conditions. But the *overall pattern* is always the same. In an anomalous situation, a routine action will either continue as usual or be interrupted. Toere is no way of *adapting* the action pattem to the situation. As examples of routine actions I have mentioned the tropistic behaviour of the flatworm, the mosquito's stinging, egg laying in the wasp *Sphex* and the well-leamed moming exercise program. Some further examples based primarily on behaviour in humans and higher primates are grasping an object, climbing a tree in a particular way, throwing something, hitting something with a

---

13 Toere are several theories of perception dealing primarily with what I call lower order perception that also present perception as hierarchical. One of the most famous is that of David Marr, concerned primarily with object recognition. His analysis and his three levels are more fine-grained than the description I give here. However, it is not essential to my arguments that the details of the hierarchical mode! I present are right. The number of levels, for instance, is not crucial.

14 In passing, it can be noted that I in fact doubt the appropriateness of Powers' third and fourth levels where object recognition is placed on level three but recognition of motion of objects on level four. Thus, perception of rigid objets is taken as more basic than perception of non-rigid movement. However, there is some behavioural (ethological) evidence which shows that recognition of motion is more basic than recognition even of particular shapes. (And from an evolutionary perspective this is not implausible.) It is also possible that some recognition of *sensations* (level-2-perceptions) can be based upon the recognition of motion and movements. (Cf. Runeson, 1983.) In any case, my discussion of perception and of planning will not binge upon whether the *details* of the hierarchical model of perception that I sketch are correct.

stone, strippinga leave from a branch, walking, jumping, stretching out an arm, bouncing a ball, and so on.15

Finally, a sixth-order perception is a perception of *a program action or program event.* These nations were also introduced in the foregoing chapter. A program action is an action regulated by a program. A program was defined as a representation of some final or overall goal and a structure of action instructions with 'branching points' where the instructions can be either in the form of specifications of subgoals or in the form of specific action representations. As stated earlier, a program can also have parts that are programs - subprograms. Programs can be more or less abstract or general, and more or less encompassing. In this way, the sixth order can be split up in a hierarchical structure of its own.

4. Relations between Perceptions of Different Orders as Relations of Invariance

In this section I want to argue that the relations between perceptions on the different levels are *invariance relations.* First-order-perceptions, that is, perceptions of intensities, are produced as direct results of physical phenomena just outside the system, such as electrical currents, light, chemical concentrations, flows of heat, mechanical influences, touch, etc. These produce activity in sensory receptors in the retina, skin, ears, etc. What is *perceived at this leve/* is nothing but *intensity.* In *one* first-order-perception there is no information to identify whether this intensity is from an electical current, from a chemical substance or from pressure. This identification does not occur until the next level, the level of sensation or quality perceptions. The incoming first-order-perceptions - and thus the physical events - can *vary* in certain ways without changing the second-order-perception. In this way, more *global* phenomena, such as temperature or pressure in contrast to local flow of heat or local deformation, can be perceived and represented. (See Powers, pp.36-37, pp.101-105.)

The same principle applies further upwards in the hierarchy. Just as a second-order perception represents an *invariant* relative to the first order (local intensities may vary, while pressure or temperature remain invariant or constant) third-order perceptions represent invariants in relation to level two. Colour, lines, edges, pressure, temperature, etc., may vary while forms, distances and objects remain constant. The configuration perception will change only if the quality-perceptions on which it is based change in certain ways. Fourth-order perceptions of speed, rapprochment, etc., are invariant in relation to certain sets of third-order or configuration-perceptions. A fifth order perception is, in turn, an invariant function of a set of fourth order perceptions. Perceptions of routine actions or events such as 'running', 'jumping', 'grasping an object', 'rolling down the slope', 'bouncing a ball', etc., are independent of the exact composition of particular movements. As long as there is an overall structure; a sequence of some particular kinds of

---

15 A routine *event* in contrast to a routine action is a stereotype process or pattern of change (in particular, a pattem of motion) that is not a change in a living organism. Examples: an object falling from a tree onto the ground, something rolling down a slope, a ball bouncing up-and-down, etc. A routine event is tied to specific circumstances, in particular to specific starting conditions. Routine events as well as routine actions have a high degree of predictability, both as to when they will occur and how they will unfold.

movements and some end condition (either the execution of an end movement or the arrival at an end situation), and this is perceived, there is a perception of *this* routine action or event.

Finally, we have sixth order perceptions, that is, perceptions of program action. A program action is invariant over certain 'subprogram actions' and routine actions. A program - and a corresponding program action - is, I have proposed, defined by a representation of a final goal anda structure of action instructions with branchings (altematives). Where there is a *branching point,* the path taken will vary between particular program executions or actions, and also the path taken toa *specified subgoal* can vary between different program executions. (While all are yet executions of the *same program* and are instances of the same kind of program action.)

The program is the whole structure of goals, subgoals, subprograms and routines and is *invariant* over the different ways of passing through the structure, both over given alternative ways and over bow to get to a specified subgoal. In general, a more abstract and composite program (and corresponding actions) is invariant over 'the ways to fill in the details', over certain subprograms and routines (and the corresponding actions). (For instance, a program for berry-gathering can be invariant over a program for gathering Red-Berries and a program for gathering Plumb-Berries, and over routines for say throwing a stick and shakinga branch.) A particular *program action* is *perceivable* as long as the specified goal, subgoals and subprograms are realized.

I want to make some further comments on what it means to *perceive* a program action and a routine action respectively. To perceive and recognize program actions means to perceive and recognize larger and more abstract (general) behaviour units than routine actions. This requires an ability to abstract from details from the level of routines and upwards. A program-perceiver might, for instance, perceive an invariant 'pick berries' in different routines or programs for picking black currants, picking raspberries and picking strawberries, etc., or an invariant 'maving somewhere' in routines for car driving, bicycling and walking or an invariant 'cleaning an object' in a certain set of techniques or routines. More importantly, it requires an ability to perceive means-end-relationships and other relationships that can tie separate routines or subprograms together in the end by relating them to some overall goal. For instance, perceiving 'someone looking for something' in a person's walking into different rooms, lifting objects and looking under them, bending, opening doors, sitting down for some time scratching his head, etc.

Consider some more examples. A system that can only perceive routine actions could perceive a situation as a group of routine actions and events, like 'I am on -abus; this is a raute that I don't know; I have never done this before', whereas a program action perceiver might perceive the situation as 'I'm on my way into town (as I am usually at this time of the day) but going by bus instead of bike and takinga different raute than usual.' Or when a routine action perceiver perceives a situation as 'Somebody is trying to open a door; the door does not open; he throws stanes at a wall and a window; he looks at the window', a program perceiver might perceive it as 'Somebody is locked out and is trying to get someone to notice him by throwing stanes at the window, so that he can be let in.'

It is not the case that the program perceiver *does not* perceive smaller behaviour elements or details - routines, movements, etc. - but that it can and does *as well* abstract from detalls and relate different routines to one another.

Some extraction of *invariants* seems to be absolutely necessary for any system that is to behave sensibly in an environment. This is so because even for kinds of behaviour that are 'hardwired' and stereotype there is little sense in relating the behaviour to matrixes of intensities. It is not proximal energy that is interesting for organisms but certain proximal patterns that can relate to the distal layout.

One theorist who frequently uses the term 'invariant' is J. Gibson (1979). He speaks of *invariants in the light* (in the ambient optical array) that *specify* objects, events and properties in the environment of the organism, including the dispositions and motions of those objects: surfaces, paths, obstacles, support, stability, solids, fluids, places, and so on. For invariants that are significant for a particular oganism or species, Gibson introduces the term *affordance.* As examples, one can mention the ground's invariant of solidity that *affords* walking on for humans, whereas its invariant of friability *affords* burrowing behaviour for moles and worms. A wool slipper may afford warmth of foot for a person, gum stimulation fora teething puppy and nourishment fora larval moth. (See Gibson, 1979; cf. also von Uexktill, 1970.)

However, Gibson's view is that these invariants can be '*directly picked up'* from the flux of information available in the organism's *sens0ly* arrays. He avoids or even rejects any use of terms such as intemal (or mental) representation, process or determinant. It is not altogether clear to me whether he is saying that this is *not what we sha/l investigate,* as the task of a theory of perception is to try to analyse the behaviour of organisms and their environments and thus to specify what organisms perceive, or whether he is saying that no intemal determinants or processes are actually *required* for producing perception. If the latter is his position then I believe he is wrong. Let us agree that there is an immense amount of information about the structure of the world that can be 'picked up' from the properties of surfaces and the way in which they reflect and transform environmental energy as this is passed to the organism's sensory receptors. Yet it is difficult to see how this information can be *picked up* without there being same *inference* going on, and, furthermore, to see what can carry out these inferences from the environmental energies to the distal objcects which transformed them if we do not assume same *internal* mechanisms within the brains of organisms.

Someone who, in my mind, has more to say on why and how invariants are perceived is R. Shepard (1984). His basic idea is that perception is guided by *internalizations of* certain, more or less enduring, *invariants or constraints (regularities) in the external world.* He discusses, for instance, the constraints of kinematic geometry which - intemalized- govern our perception of motions of rigid objects or of local parts of nonrigid objects during brief moments of time.

Shephard presents a classification of what he calls *internal determinants* of perception. These determinants are not available in the optic array or any corresponding arrays of the other senses at the time of perception but are internalizations of current or previously prevailing extemal circumstances. The classification is by the remoteness of this origin in external circumstances; going from determinants *temporarily* established by the current context (like attentional biases established by just foregoing and ongoing perception, and in this sense by preceding external circumstances); over determinants aquired through past experience where certain constraints have been internalized through learning or perceptual differentiation (like chess masters that discriminate hetter between board postions that might occur in an actual game of chess than between ones arranged at random); to detenninants incorporated into the genetic code during the evolution of the species. It is in the latter that one, according to Shephard, would expect to find internalizations of the most enduring and ubiquitous invariances in the world. *(Ibid.,* pp. 431-432.)

Such intemal determinants are in my terminology carried by *perceptualfunctions* and *recorded values* of such functions (that is, recorded perceptions). A long term determinants can *be* a particular perceptual function or recorded perception, or an aspect of such, and a short term determinants can be conceived of as the *activation* of a particular perceptual function ora recorded perception. The activation of a function as well as the acti:vation of some expected values in the form of activated recorded perceptions prepares for the handling of certain information and may influence perceptual processing. In this way there can be a direction of attention and a filling in of 'missing' but expected perceptions and so on. I will shortly retum to this.

5. Perception as Subjective I Constructivistic

The theory of perception that I present is a *constructivistic* theory of perception (cf. pp.16-17). What a system perceives -the system's perceptual environment-depends on *internal determinants* as well as *external determinants,* that is, on the system's extemal environment; on what is there and happens and directly affects the system's sensory arrays. Internal determinants influence perception by selecting and synthesizing information and by disregarding and adding information. Consider, for instance, skin perception. This is a result of a combination of perceptions of temperature, pressure, surface structure, etc. It is a *unitary* perception but we could think of another perceptual function that creates a skin perception in a slightly different way, so that, say, temperature would not be involved or that pressure would not be one of the features or that some other sensation or dimension would be involved.

Or consider the perceptions of 'somebody running', 'somebody waving a hand' and 'a bus at a bus-stop' compared to the perception of 'somebody in a hurry trying to make the bus wait by waving her hand to the bus driver'. The latter perception, a perception of a program action, is a synthesis of lower-order perceptions and could not have occurred without some intemal determinants - some intemalized constraints or regularities - in the perceiver.

Or take the phenomenon of apparent motion, where the same altemating visual or auditory display can lead to distinctly different apparent motions. Indeed, if one presents two differing input-arrays of configurations successively toa human being, one out of three perceptions may arise. One may see an object disappearing and being replaced by another one, or one may see one and the same object moving or one may see one and the same object changing in shape so as to be transformed into something different. (See Boden, 1983, pp.18-19.) Determining which of these perceptions actually occurs must be due to some intemal factors. Thus, perception cannot adequately be described as simply picking up invariants 'completley available' in particular extemal stimuli. What kinds of information are synthesized and bow, what is focused on, added, neglected, etc. - the invariants perceived - depends on the perceptual functions in question, for instance on which perceptions from the level below that these functions take as arguments and how they process them.

Powers is explicit about the constructivistic element in his ideas. He says that perhaps we ought to speak of perceptual functions constructing rather than recognizing perceptions; for instance speak of *"sensation-creating* input functions rather than sensation-recognizing functions". (Powers, p.113-114.) To use the term *recognition* implies that the environment contains one particular order to be recognized, and that all we have to do is learn to detect this one order. Furthermore, in his view, there is nothing in principle to prevent a nervous system from performing computations leading to perceptual signals that have no external significance. "[P]erceptual signals *depend on* physical events, but what they represent *does not necessarily have any physical significance."* (Powers, p.37; pp.113-114.)

Perception is not pure *bottom-up* processing, that is, processing where what takes place on higher levels does not influence what happens on lower levels. There are elements of *top-down-processing.* Perceptual activity is influenced by *anticipations* and *expectations* of at least three kinds:

• Expectations of, and preparedness for taking in, certain (kinds of) information, so that this particular information rather than other information is *attended to;*

• Anticipations of, and preparedness for taking in, certain (kinds of) information, so that the perceiving system *adds* or *fills in* such expected information that is missing, and *neglects* or *dispels* information that does not fit in with the expectations;

• Expectations and anticipations that set off *active search* for particular (kinds of) information. [16]

An example: A hungry animal searching for blueberries will *attend to* something blue rather than attend to colour nuances and details in the branches that lie on the ground. And in seeing something blue, it may indeed have a perception of blueberries - round shapes, smooth shiny surfaces, something with volume - even if the blue colour comes from a blue scribbled paper. The organism is also *searching,* maving around, touching, looking for blueberries and not just sitting passively 'getting to see and feel' things.

Expectations can be seen as carried by perceptual functions and recorded perceptions (cf. the discussion about internal determinants, pp.76-77). An expectation can either *be* a particular perceptual function or recorded perception or the *activation* of a particular perceptual function or a recorded perception. Perceptual functions and perceptual values, and their activation, imply preparedness to take in and process certain information (lower-order perceptions), to attend to certain information, to fill in certain information, etc.

In almost *all* perception there is some of these anticipatory moments. Perception is always *one step ahead of itself,* filling in expected information that then may or may not also be supplied 'from the bottom'. This is not paradoxical if perception is viewed as a continuous activity where activated information-structures or representations direct perceptual activity which then modifies these representations, and so on. Thus, in this model, *lower* does not mean *earlier* and *higher* does not mean *later.* Perception -like behaviour-is a

[16] All these, but in particular the last, play a role in appetence-behaviour (see chapter 3, page 23).

continuous process. The perceiving system is always in some context, perceiving and behaving, and has consequently a particular preparedness and certain expectations. Perception at one moment of time is to a great extent determined by currently ongoing or preceding perceptions and behaviour in the system. It is not a question of dealing with discrete stimuli.

In sum, perception is *subjective* in the sense that it is the constitution and behaviour of the system, as much as physical events outside, that *determines* what the system in fact perceives. *First,* it is the transduced information *together with* determinants *inside* the system - information-structures, functions and processes - that result in perception. (Even the constitution of the sensory receptors implies *afirst selection* already on the level of transduction. These receptors are sensitive for *particular* light-frequencies, certain pressure-frequencies, etc.) *Second,* the organism influences what it perceives when it *behaves,* that is, when it moves, turns its head, maves its eyes, ears, touches, etc. Organisms do not just sit passively and wait for their sensory receptors to receive input. They are, as I have pointed out, not purely re-acting, but also displaying activity of their own, activity determined from within. They display appetence behaviour. (Cf. pp.16-17, 22-23.)

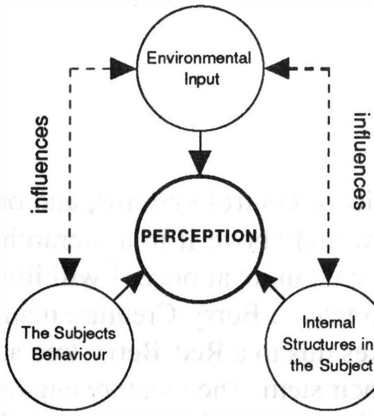


Figure 11. The subjectivity of perception

6. Behaviour as the Contra/ ofPerception

Power's main thesis is that behaviour is *the contra/ ofperception.* The idea of control mechanisms and feedback involved in behaviour, as such, is not astonishing, original or revolutionary. Some of the subtle interactions between perceptual and motor systems in living organisms are comparatively well studied. For instance, when stretching out a hand to retrieve an object there is feedback from each particular movement. There is a comparison of the goal state and the actual position of the hand at any particular moment in time, and, on the basis of this, subsequent movements are regulated.

Indeed relatively much is known about feedback mechanisms in the operation of the bodily kinesthetic systems in human beings.What Powers does is to *extend* this kind of analysis. Apart from the bodily kinesthetic system, he also considers control of higher-level behaviour or action, where not so much is known about the underlying neural mechanisms, and he attempts to relate different kinds of control or feedback systems to one another. My aim in this and next section is to give an account of Power's analysis.

The basic control unit of a perceiver-and-behaver is presented in figure 12 (adapted from Powers, p.209). Its basic principles are the following. A *reference perception* or a 'desired perception' is compared to the *present perception.* More precisely, the *present value* and a *reference value* of some perceptual dimension are compared. The difference between these (the error) is the basis for some kind of action instructions for obtaining the reference value of this perceptual dimension. These instructions, we will see, take the form of reference perceptions for lower levels.

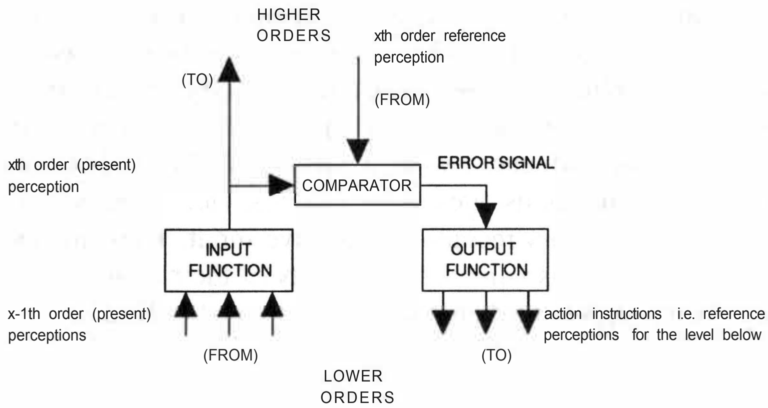


Figure 12. A perceptual-behavioural control system

On each *control level,* there is a *huge number* of such units or control systems, and one particular control system relates to control systems below and above it in a hierarchy. Behaviour as a whole results from the operation of *many* such units at once. I will illustrate the basic principles by considering the following example. A Berry-Creature manufactures a tool by stripping off leaves from a branch. It takes this to a Red-Berry-tree and uses it to hit ripe Red-Berries, aiming at and trying to hit their stem. The creature engages in this until it has got a couple of berries on the ground, and then it climbs down to eat the berries. It is an execution of its program for getting Red-Berries. In executing this program, the creature adapts to various environmental factors such as the following: the ripeness of the berries, where the particular berries are placed (how high up, etc.), the wind, properties of the tool (thickness, suppleness, slipperiness), etc.

But how is the program executed? Let us approach this question by considering what the organism *perceives* and what it *does* 'on different levels'. The central question is: 'what perceptual dimensions are perceived and controlled?' Let us concentrate on the creature's aiming at and hitting a Red-Berry stem with the stick. This is a common kind of behaviour in the Berry-Creature's repertoire; a routine action with a fixed overall structure. Toere is no greater variation in the particular movements of the creature's limbs or in their sequencing, nor in the end condition: to have the stick hit a berry. But at the same time, it is evident that the *details* of particular movements (approaching, retaining, the speed of particular movements, maving right, maving left a little bit, etc.) cannot be fixed and exactly the same every time the creature is to hit a Red-Berry. There must be same flexibility in response to particular environmental conditions (wind, berries position, properties of the stick, etc.) if this behaviour is to be successful. When the creature aims at a berry, there is a goal or reference perception to be obtained, namely, the perception of the stick hitting the berry (to see and feel this), which is the end condition of the routine.

What particular movements - and thus what reference values for movement dimensions - are required moment by moment depends in the first place upon *the routine* that is executed (in other words, the routine reference that is acted upon) but also on environmental conditions. And, in turn, in relation to those *movement reference perceptions* and environmental conditions, *particular configuration perceptions* are required or 'desired'. Finally, *certain perceptions of sensations* are required in order to obtain or maintain a particular configuration perception.

On *each level* there are *several systems* that *control* their particular perceptual dimensions: various sensations, positions, objects, movements, routine actions and events, etc. That is, they obtain or maintain particular values of those dimensions (like a particular taste, a particular position, a particular stick movement, etc.) in spite of disturbances in the environment and varying conditions between the different occasions when such a reference value is to be obtained or maintained. In order to maintain, for instance, a perception of a stick straight in front of the berry stem that one is aiming at when the target moves because it is windy, the reference perception for *hand motion* must be changed. Now, to obtain this, the reference perception for *hand position* must be changed. And in order for the hand to be kept in any one position in spite of any mechanical disturbances, reference perceptions for *muscle tensions (sensations)* must be changed. (See Powers, pp.51-55.) Furthermore, if certain muscle tensions and efforts are normally required to keep a stick in a particular position in one's hand, but the stick is unusually slippery that time, we must have other reference perceptions for muscle tensions.

The reference perceptions on one level are *generated by* the level above. The *behaviour instructions* that a higher control level gives to lower levels are namely in the form of reference perceptions or values. As an example, a fourth order system will compare a *motion* reference and a present motion perception and use this comparison as the basis for generating new or modified reference perceptions of *configurations* for some third order system(s). In the end, it is *muscles* that are the effectors. *Sensations* - second-order perceptions - can be directly controlled by combinations of *muscle tensions.*

Thus, on several levels there is adjustment to disturbances in the environment that influence controlled perceptual dimensions (the wind that is blowing, resistance, objects moving, etc.) and to varying conditions between different occasions when a system is controlling a particular perception.

What we have considered so far is how behaviour up to the level of routine actions can be controlled: how a particular routine action can be brought about and be adapted to the particular environmental conditions. However, such behaviour only involves what I have called *local flexibility.* In routine behaviour the overall behaviour pattern is fixed. There is flexibility with regard to details but within rigid frames and only involving continuous and quick adjustments to local conditions. But Powers goes further. He claims that the principles for behaviour control[17] on the level of *program actions* - i.e. for producing

---

[17]***Behaviour*** as such is never ***controlled.*** Only perceptual dimensions - ***and percevied behaviour-Can*** be controlled. When I speak of controlling a routine action or controlling a body position, this is short for controlling a perception of a routine action or a perception of a body position. Also when I speak of retrieving or generating ***behaviour instructions*** or ***behaviour references,*** this means retreival or generating of ***reference perceptions*** of positions, movements, routines, etc. The reason why it is still appropriate to

program behaviour - are just *the same* as for lower-order behaviour. Toere is a reference perception - now of a program action - to control; that is, to obtain or produce as a current perception. And on the basis of the difference between the reference and the present time program perception, references or instructions for subprograms or routines are given as output. And this continues down the hierarchy, until the outputs are muscle instructions. It is particularly important that control of program actions involves control of (perceptions of) *means-end-relationships.* Toere is *control of subgoals as related to higher-order goals.* That is, the system strives to obtain or maintain *such relations* in spite of disturbances and variance in environmental conditions.

So, the output of a control system on one level is *reference perceptions* for lower orders. But *what are* reference perceptions? Where do they come from? A reference perception, according to Powers, is a *retrieved recording of a past perception.* It comes from memory.[18] (Powers, p.217.) To be precise, the output from a system on one level does not consist directly of reference perceptions for the level below as such but of memory addresses for memory systems on this level. See figure 13 (adapted from Powers, p.218.). And then there is a selection from memory of the values of particular perceptual dimensions that are to be recreated in present time: a particular muscle tension, a particular body position, a particular object, a particular routine action, etc. Memory, just as perception and behaviour, is hierarchically organized, and every control system has its own memory. (Powers, p.208.)

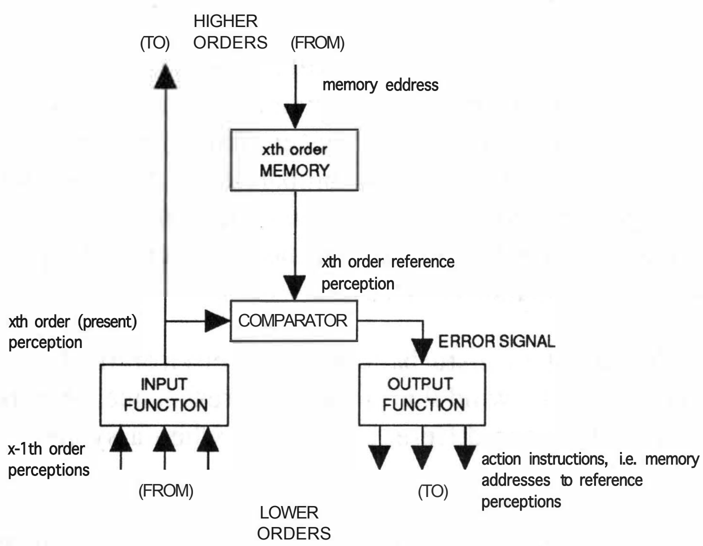


Figure 13. Perceptual-behavioural control system unit with memory

This occurrence of reference perceptions on different levels implies that, parallel to the hierarchy of perception and the hierarchy of behaviour (see p.72), we have a hierarchy of

speak of behaviour instructions is that the retrieval of a reference perception in general means immediate acting upon this. For instance, in order to reproduce, say, a routine reference perception, the system will act and try to execute the corresponding routine.

18 And so reference perceptions are ***representations*** in two ways: first, because also a ***present time*** perception is a representation in that it ***organizes*** and presents information ***in a particular format,*** and second, because by being retrieved from ***memory,*** a reference perception ***re-presents*** a past perception.

goals. The *goals* are to bring about or maintain these reference perceptions as present time perceptions. [19]

behaviour	*perception*	*goals (reference perceptions)*
program action	program perception	overall goal + subgoals
routine action	routine perception	end condition + overall action pattern
movement	motion perception	a particular movement
limb position	configuration perception	a particular configuration
muscle tension	sensation perception	a particular sensation

Figure 14. Behaviour, perception and goals

How does this relate to what I said about goals and goal-directeness in section 3.3 and to the biological goal or interest hierarchy presented there? A biological goal hierarchy is a description of the complete behavioural sphere of an organism and of all interests with corresponding goals of an organism. It is also a description of how lower-order goals and interests relate to higher-order goals and interests and in the end to the *primary biological goals* (of self-maintanence and self-propagation) which are goals by definition. A hierarchy of reference perceptions as the one in figure 14, on the other hand, is a description of the performance of one particular kind of behaviour. But it is possible to take a hierarchy as in figure and fit it in somewhere under some interest in figure 1, p.24. The relations of goal and subgoal, or goal and derivative goal, will then apply. However, the *original analysis* behind the goal notion in the respective hierarchies is not the same. A biological goal is a goal because of its relation to primary biological interests and goals. A goal in a perceptual hierarchy is a goal if it is the case that disturbances tending to cause a deviation from this perception call forth a behaviour which results in counteraction to the disturbance.

A final feature to note is that Powers claims that his theory of behaviour is in a sense a *complete* and *general* theory. Posing the question of when his theory is applicable, he indeed proposes *"all of the time."* (Powers, p.47.) His main claim is "[that] all behaviour is oriented all of the time around the control of certain quantities with respect to specific reference conditions. The *only* reason for which any higher oganisms acts is to counteract the effects of disturbances (constant or varying) on controlled quantities it senses." (Powers, p.47., my italics.)

In my mind there is no reason to go this far, and there is indeed evidence against such an encompassing claim. Think of certain highly skilled 'automatic' actions of, say, the athlete, the pianist, the typist. Some of this behaviour - long sequences of movements - proceeds so rapidly that feedback from perceptual and kinesthetic systems can hardly be used. The actions unfold at such a great speed that modification in the light of perceptual information is hardly possible. Or think of other 'pre-:programmed' sequences that are not

19Powers defines the *goal* of some behaviour as the reference condition of the perceptual dimension that is controlled by this behaviour. And the *purpose* of the behaviour, he says, is to prevent the controlled perception from changing away from this reference condition or value - to maintain it, or bring it about. (Powers, p.50.)

learned but innately specified. E. von Holst (1969) has shown for instance that in certain species there are certain sequences of movements, used in locomotion, that depend only on central systems and also unfold when all sensory information channels are cut off. Yet another example is Lorenz' 'Leerlauf-Bewegungen' (cf. p.48), which can be very complex motion patterns, that unfold without sensitivity to perceptual information. (Lorenz, 1973.) To me there seems to be reason to retain the likelihood of routine behaviour without feedback and maybe also certain kinds of interactions with the environment, based on a moment-to-moment responsiveness, which do not involve feedback and yet produce adaptive behaviour.

The point I want to make is the following. I do not believe that all explanations of adaptive or goal-directed behaviour - whether in lower or higher organisms - is to be found in Powers' kind of analysis or in any 'feedback analysis'. However, I *do believe* that this kind of analysis can account for *many aspects* of a *main part* of an organism's behaviour, in particular aspects of what I have called intentional behaviour (behaviour that is initiated and maybe guided by some goal representation). I furthermore believe that it is *in this kind* of behaviour control that one can look for the origins of a planning capacity.

7. *Perceptual-Behavioural Systems of Different Orders*

In this section I address what Powers calls perceptual-behavioural systems of *different orders*. It is perceptual-behavioural systems of *higher orders* that are most interesting if one wants to develop a system capable of planning. Here I also introduce the concept of *trial-and-error-behaviour* which is going to be useful.

Two perceptual-behavioural systems of different orders differ in the following respects:

• What kinds of invariants it can perceive and control - the highest order perception that the system can perceive and in relation to which it can generate lower-order references

• What kinds of disturbances it can perceive and counteract

• What kinds of perceptions and behaviour it can relate to one another and coordinate

• What kinds of behaviour (elements) that are accessible, separately or independently, to the system.

• Whether it can learn, and what it can learn: what kinds of behaviour can be *tried out*- and maybe learned- and what has to be genetically determined and/or determined by a particular situation (a system's perceptual environment and its motivational state).

These points will be illustrated by some exampels. Let us start with some perceiver-and-behavers of low orders. In one example where Powers discusses control of arm-hand movements and positions, he contrasts a *second order* and a *third order* perceptual-behavioural system, i.e., a *sensation-governed* and a *configuration governed* perceptual-behavioural system respectively.

The *sensation-governed* perceiver-and-behaver cannot perceive or control any perceptions higher than sensations. Hence it cannot, for instance, control a particular *arm-hand position* in relation to environmental influences. In an arm and hand there are, according to Powers, some twenty-seven independent motions: different flexions at joints, rotations in ball-and-socket joints, etc. Together, such motions, of course, have the effect of placing the hand-arm combination in specific orientations and positions in physical space. But in a sensation-govemed system, this is not something that can be accomplished under feedback control. Second-order kinesthetic systems dealing with hand-arm-movements are only capable of controlling sensation-dimensions, like effort or tension in tendons, museles or joints. One such system can control *one such* 'direction' in this twenty-seven-dimensional 'kinesthetic space'. But this is independent of what the other (about twenty-six) second-order systems are <loing. *Arm position in conventional three-dimensional space* is, according to Powers, a construct of a configuration governed perceiver-and-behaver. The sensation-govemed system cannot synthesize lower-order perceptions to perceive a particular position. Hence it cannot detect *disturbances* that tend to cause a deviation in a position and it cannot *coordinate* lower-order behaviour (such as the independent motions mentioned in the example) to counteract such a deviation. (Powers, p.107.)

A third order or *configuration governed* perceiver-and-behaver can, however, detect such disturbances. It can control - obtain, or maintain - a perception of a particular configuration in spite of disturbances and variations in environmental conditions. The discrepance between a reference and a current perception for a particular configuration results in a generation of reference perceptions of sensations by addressing sensation memory systems. And then, discrepances between these and current sensation perceptions result in muscle tension references - or instructions. Thus the system acts in order to reproduce these different reference perceptions - and in the end the configuration reference - as present time perceptions. A configuration governed perceiver-and-behaver can perceive and control its own position and bodily configuration. A fourth order or *movement-governed* system is capable of more. Not only can it control, for instance, particular positions but can also produce controlled *movemens.* It can control smooth transitions from one configuration to another and produce such configuration transitions at different and specific rates.[20] It can produce a particular movement in spite of environmental disturbances.

However, with a movement-governed perceiver-and-behaver, the *movement references* must, according to Powers (p.137.) originate either from immediate environmental conditions or from genetically transmitted information. In my conception, they have to be determined by the system's current perceptual environment and motivational state. Every time the system is in a certain perceptual environment and a certain motivational state, the same movement references are given or determined. The individual cannot *vary* and control movements in relation to some higher-order behaviour, like a complete routine action. Now, such a system is very primitive and not of much interest if we want a system that can be developed into a planner. So leaving the lower orders of perceiver-and-behavers, I will now dwell more on routine-governed and, in particular, on program-govemed perceiving-and-behaving systems.

The *routine-governed* system can perceive and control routine actions and events. A routine reference perception is compared to the current routine perception, and, on the basis of this, addresses are sent out to retrieve particular movement references. In the same

---

[20] 'fhis is an aspect where fourth-order or motion-control differs from third-order or configuration-control. In the latter, all transitions, which are induced by a suddenly appearing configuration reference perception, occur ***immediately*** - as fast as the control systems are capable of acting. (Powers, p.135.)

way, position references and, in the end, sensation references are retrieved. The generating of and acting upon these references proceeds in order to reproduce the routine reference as a present time perception; that is, to reproduce a perceptual invariant of a *particular movement sequence* and an *end condition* (either an end movement or an end situation). What drives the behaviour of the system - what it strives at - is to reproduce the particular routine perception.[21] When it has gone through the movement sequence and reached the end condition, either by executing the end movement or reaching the end situation, it will come to an end. Now, consider the shortcomings of a routine-govemed system. Such a system, as it is not capable of perceiving and controlling any higher-order behaviour or events than routine actions or events, perceives each routine action and event only as separate and not as related to any other routine actions or events. Consequently it has no perception or representation of a structure of *subgoals,* where lower-order goals are related to higher-order goals and in the end to an overall goal; that is, it has no capacity to perceive a structure that *ties together different routines (or subprograms).* Thus it is also not capable of *controlling* a routine action - of performing a routine - *in relation to a program reference.* When the environmental conditions diverge from those to which a routine action is adapted, the routine-governed perceiver-and-behaver gets into trouble.

Asssume, for instance, that a creature has a routine for collecting and eating nuts, which involves letting a gathered nut fall to the ground. A routine, remember, is specified with respect to its general movement pattern. Using this routine works well under normal circumstances, but say now that a particular tree stands on a slope and near the water. To perform the routine action 'let-nut-fall' here implies that the nut rolls into the water. But the routine-governed system cannot do anything about this. It can only repeat the same thing with the next nut.

A *program-governed* perceiver-and-behaver, on the other hand, may see that it maintains or obtains a particular program perception, i.e., a *structure* of an *overall goal* and action instructions in the form of certain *subgoals* and *kinds of actions.* Thus a program-governed system in this case has some possibilities to control the *subgoal* of getting the nut to a place where it can then pick it up to crack it and finally eat it. It can counteract disturbances in this subgoal-goal-relation by varying the routines: producing another routine action than 'let-nut-fall', such as, for instance, a routine action of throwing an object away. Or if a given nut is extremely hard, the routine-governed system will only apply the routine it has for cracking these nuts. It will repeat it again and finally probably stop out of tiredness - with the end condition of the routine not yet obtained. Whereas a program-governed system again might apply some other routines - rubbing; kicking, biting, hitting with a stone, etc. - relative to the subgoal of cracking the nut. Or finally, if the *routine-governed* system arrives at a situation that is so unfamiliar that no routine is activated at all, it cannot deal with this in any sensible way. Say that a big tree has fallen down over the path and that this is a very unusual happening. A *program-governed* system might be equipped to handle this. It might, for instance, have some general program for dealing with obstacles and within this try out some subprograms or routines like climbing, biting,

[21] A routine perception represents the fact that a given routine action or event is in progress or has occurred.

creeping under, hitting, walking along, etc. And it might thus find a way of overcoming this kind of obstacle.

The example I just gave where an individual apparently *tries out* different action pattems is an example of what I call *external* or *behavioural trial-and-error.* What we have generally here is:

• A start state - where the trial-and-error behaviour starts - defined by the organism's perceptual environment and motivational state

• An end state or end condition - where the trial-and-error behaviour terminates - which is a state corresponding to the motivation and interest: either the particular motivation ceases (the motivational state is changed) as the interest is satisfied or met like when one tries to escape from a threatening situation and succeeds, or the system attains a (familiar) situation where it can get on with some routine or program behaviour relating to the interest, like when one has obtained the nut and can get on with an open-nut routine or program. (Toere can thus be more than one end condition.)

• A repertoire of behaviour elements

A *trial* consists in the execution of one or several behaviour elements. Reaching an end state means success, not reaching one means *error.* (See figure 15.)

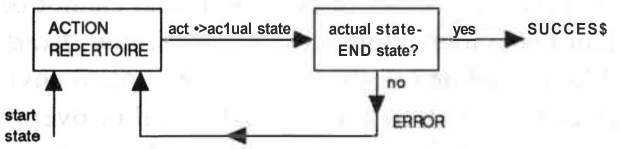


Figure 15. Behavioural trial-and-error

Behavioural trial-and-error, I claim, is something that occurs in a limited way on the level of routine control and first really on the level of program control. The reason why *lower-order* behaviour is rarely trial-and-error behaviour is the following. Behavioural units below the level of routine actions are short and swiftly performed. The behaviour controls reference perceptions that are temporarily limited, and the environmental conditions and the disturbances in relation to which the behaviour occurs exist only for a short time. A particular motion may be evaluated in the sense that it is compared to a reference. Maybe the motion was an error in that the discrepance is large. However the next moment the situation is changed: for instance, the arm is already in another position, or the object one is aiming at has moved in the wind. This means that for the lower control systems there is a new problem situation and new goals. For instance, other configuration perceptions are now the desired ones. The detailed environmental changes of relevance for lower-order behaviour are swift, and action also swiftly transforms the problem situation for lower-order control. We do not have here one relatively constant problem situation and a corresponding end state (an overall goal) in relation to which *several alternative actions* from a repertoire can be *tried out,* but there is a new problem and goal for each moment of time.

It is first on the level of program-control that there is real space for trial-and-error behaviour. Here the situation in relation to which the behaviour is performed, i.e., the end state and the environmental disturbances that are to be counteracted, are not so local in time. Several routines or subprograms can be tried out in order to obtain a goal or subgoal within the frames of a program action, because a program for action, as we know, does not have to be specified down to every routine. In controlling a program reference, an individual can engage in trial-and-error behaviour to obtain specified subgoals as in the example above, where the disturbance in the form of a tree that has fallen down, and the end state of having passed it, are relatively constant, and where different ways for getting over this obstacle can be tried out. Also in a very unfamiliar and thus 'problematic' situation where all that is immediately clear to the system is some general overall goal but no program-structure at all, it is still possible for a program-governed system to randomly try out routines and subprograms. *In general, when* a program-governed system has a primary motivation and is in a *novel* kind of situation - where novel means unusual, unfamiliar, etc. - we can expect the system to engage in behavioural trial-and-error.

A useful extension of a system capable of behavioural trial-and-error is - of course - to endow it with a *learning capability* so that new and successful behaviour patterns that it tries out can be stored in some way and used again in similar situations.[22] (Note, however, that a capacity for trial-and-error behaviour does not *imply* a capacity for learning.)

Observe that it is only possible to *learn* a behaviour pattern that one may - in some sense - *try out.* In other words, it is only possible to learn a behavioural dimension (a hand movement, a climbing routine, a locomotion program, and so on) that is 'open' and where various alternatives can be tried. But there are kinds of behaviour that cannot be tried out in this way by a system. Certain behaviour references are, we know, *fixed.* They are given or determined by the problem situation (motivation and perceptual environment) and cannot be adjusted or 'replaced' but indeed constitute the final or overall goal (now blue-berry-program, now jump-into-the-tree-routine, now move leg, now assume this position...). Action references of the *highest order that a system can control* are fixed in this way.

When learning occurs in *lower-order systems* this is primarily learning of the kind where the system learns to apply a particular fixed behaviour pattern in other, but similar, situations than the original or usual. But lower-order systems cannot really generate *new kinds of action* to suit 'novel' situations. Systems below the level of *program-control* cannot generate and learn a new composite behaviour pattern by combining (parts) of routine and program actions from its behavioural repertoire. A *routine-governed* system might indeed *perform* a new sequence of routine actions in a difficult and unfamiliar problem situation. But even if that 'trial' would be a success, the behaviour will not be learned. The system is not capable of perceiving such a composite behaviour pattern as a whole. It only handles each separate routine independently and matches this directly against the overall goal. But it cannot, as the program-governed system can, coordinate several routine actions, perceive their consequences as subgoals and perceive them all as a unit by being related to the overall goal. Assume, for instance, that we have an enclosure of a completely unfamiliar kind and that an organism is motivated to get out from this because it is hungry or tired, for example. The routine-governed system might

---

[22] By a system's ***learning*** I mean change in probability relations between a kind of situation (perceptual environment) and a kind of behaviour in the system, where this change is due to specific experiences of the system and is a change in an adaptive direction.

perform certain routine actions: bite, start kicking, raise an arm, throw something at the wall, grasp it, touch it, kick it with its leg. But it cannot perceive a consequence of such an action as a subgoal - a part of the wall cracking for instance. And it cannot coordinate routines. Maybe it would succeed if it would hit with something and kick with its leg simultaneously. It can apply its routines as wholes and may also execute a *sequence* of routines, but it cannot perceive and store such a composite behaviour unit. The program-governed perceiver-and-behaver can do these things. Furthermore, it does not just perform entire routine actions but may execute *parts·* of routines and programs in novel combinations.

We realize that a program-governed system in many ways is more powerful as to what it can perceive, control and learn. A program is more than a set of unrelated routines. This becomes apparent when the program-governed system engages in trial-and-error-behaviour. For instance, when there are some alternatives specified in a program and one does not work (i.e., it does not work in relation to the overall goal) the system may perceive this and try another alternative. A program may be interrupted and then continued later on, etc.[23]

It is a *program-governed* system, capable of trial-and-error behaviour and of learning, that I will use as a basis for a planner. Furthermore, it is a system endowed with *curiosity,* which means that it will make *use* of its capability for trial-and-error behaviour and for learning - trying things out, and possibly learning from this - also in situations where there is not a serious problem at hand. Recall that curiosity is one of the two *secondary* motivations of an 1-Creature. It relates to the interest of *seif expansion* and motivates exploratory behaviour and play. Thus a system endowed with this can use its capability for trial-and-error and learning also in play and exploration.[24] Moreover, curiosity may also influence behaviour related to primary interests and allow some trial-and-error in spite of the fact that an individual already has a suitable routine or preformed program (cf. p.29). Let us take an example. A Berry-Creature is motivated by hunger and acts upon a Red-Berry-program suitable for the current environmental situation. Under certain circumstances though; if the hunger is not too strong, and the individual is motivated by curiosity, the system might quit - interrupt or terminate - the program execution. On the way - the habitual way - to a Red-Berry-tree with the tool it has fetched, it may for instance notice a new path (say that a familiar obstacle has been removed) and take off on it, instead of following the program. Or it may notice an unfamiliar thing that smells like a Good-Berry but has a hard shell around it and start manipulating it, i.e., biting, holding, hitting, throwing, etc., although this means interrupting the Red-Berry-program execution. These behaviours *could* be successful; that is, the novel_ path might be a shortcut, and the scented thing might be edible. Curiosity means an increased sensitivity to *unfamiliar aspects* of the environment and not only to what is anticipated within the frames of a program, and there may occur an interplay between curiosity and primary motivations so

---

23 Within the group of program-controlling systems, there may be great differences as to how composite and general the programs that the system can handle can be. It is an advantage to be able to abstract and generalize - as long as this capacity is complemented with a capacity to retreive and handle details.

24 I shall retum to play and exploration in chapter 10, where I will also discuss the particular features of 'trial-and-error-behaviour' in these contexts.

that, as in the example just given, the sensitivity to opportunities relating to the interest of food is particularly increased.

8. Conclusion

What we have obtained so far is a basic design of a perceiving-and-behaving system. Its basis is that such a system *behaves* all the time in order to control reference-perceptions on various levels (even if one can admit behaviour that has other grounds, cf. pp.83-84). We have discussed various aspects of this design that will turn out to be useful when we want to endow the system with a capacity for planning - aspects such as the hierarchical structure of perception and behaviour, the construction of and dealing with abstract representations (perceptual invariants), the subjective and constructive aspects of perception, the acting upon goals, the anticipatory moment of perception, and so on.

Furthermore, by varying the number of levels in the basic design, we get perceptual-behavioural systems of different orders with varying capabilities to handle perceptions and behaviour units: configurations, movements, routine actions, program actions (of varying complexity and generality).[25]

Perceptual references of the highest order for the perceiver-and-behaver are fixed, i.e., determined by the system's current perceptual environment and motivational state. Lower-order references on the other hand are flexibly replaced and modified during action in order to realize the highest order reference as a current perception. The activation of a reference-perception brings about *immediate acting upon the reference.*

There are continuous comparisons between current perceptions and reference perceptions, and, on the basis of this, the perceptual references on the level below are modified. In *normal situations* (i.e., habitual or familiar situations) the coordination between control systems on the different levels is smooth. Higher-order systems continuously and easily generate or select instructions for systems on the level below by addressing memory on that level, on the basis of divergences. For example, a program reference produces immediate retrieval of routine instructions, movement instructions, position instructions and muscle instructions. And acting upon these, in normal situations, result in adequate behaviour. In *problem situations,* on the other hand, the perceiver-and-behaver might more randomly- and maybe not so smoothly-retrieve references and act upon those. On the level of program control (and to a limited extent in routine control) we obtain *trial-and-error behaviour* in these situations.

We ended up with such a program-governed (a sixth-order) system capable of trial-and-error behaviour and learning and equipped with a curiosity motivation. This is the *basic* design of an 1-Creature. It controls certain perceptual dimensions, it acts upon goals and reaches goals. It is a kind of system that can manage quite well and produce much *planful*

---

[25] In saying that a perceiver-and-behaver is of a certain order, one only indicates the most advanced form of behavioural control in the system. It is not the case, for instance, that all behaviour of a program-governed system has to relate to a program reference. It can also at times act upon a separate given routine and execute separate routine actions. Program-governed systems may - and often do - act upon routines, habits, and reflexes as well. Note further that the difference between perceptual-behavioural systems of different orders concerns what the systems perceive and control in a situation and that it primarily is a difference from the system's points of view.

behaviour. But the system never *plans:* it is never the case that it represents a goal and a start situation and generates and tests (i.e., mentally dissents or assents to) representations of sets of action instructions for getting between the start and the goal. How to obtain this is the topic of next chapter. And this also means taking another step toward the description of the design of an I-Creature.

Illustration 4. A planner

7. How TO OBTAIN PLANNING IN A PERCEPTUAL-BEHAVIOURAL SYSTEM

1. Introduction

The question to be addressed in this chapter is the following. How can the basic design of a perceiving-and-behaving system be developed into a system that is capable of making plans for its actions? I will, in this chapter, discuss such an extension, where the capacity for *planning* is a capacity for *imaginative generating-and-testing.* Toere is first a section on why the *basic* perceptual-behavioural system is incapable of planning. Toen I introduce imagination and the idea of intemal generating-and-testing. Finally there is a section on the general use of generate-and-test models for psychological phenomena. One point of importance that I attempt to stress in this chapter is that planning may occur at the expense of other activities; that it competes for resources with behavioural and perceptual activities.

In this chapter and the next I shall use the following four examples, which are examples of planning in a Berry-Creature of the first generation, that is, of an 1-Creature. In these examples, I first describe the planning situation: the individual's motivational state, M(primary motivation; secondary motivation), aspects of its perceptual environment, PE, and some relevant features of the situation. Thereafter, I speak of how the individual is constructing a plan for action.

Example (I.I): 'Get something to eat -Red-Berries growing high on thin branches':

M(hunger; -) PE(a Red-Berry-tree full of Red-Berries on branches unusually high up)

The individual *constructs a plan* in the following way to get these berries: If I take my usual routine and jump into the tree with a stick taken from a Lind-bush in order to knock down the berries, this will not work because the stick is too short. What about trying to climb further up the tree? No, the branches get too thin fairly soon and may break- and so I would risk falling down and getting hurt. Another alternative is to take a stone and throw it at the berries (as I throw stones at Plumb-Berries to have them fall down), aiming at the shafts of a bunch of berries... it is difficult to aim – If I miss-- I must climb down and fetch the stone again. I also run the risk of crushing the berries with the stone. Use a *longer* stick then? A stick from a Long-Twig-bush? Butonethat is not so pliable, but thicker than those that I use for the Umbrella-trees. There is a Long-Twig-bush just down the slope. Yes, I will go there and get a stick..

Example (/2): 'Getto a shelteredplace'

M(freezing; -), PE(the western part of the area near the river, at A (see map below), strong wind blowing)

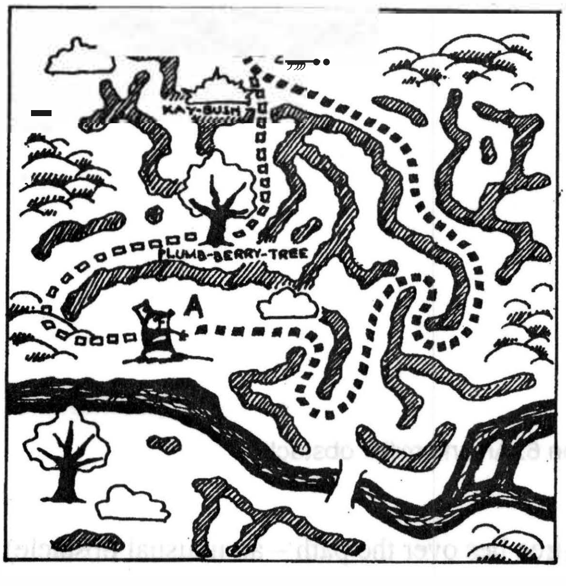


Illustration 5. How to get some shelter

The last couple of times the individual has been is in this area - exploring and playing - and it has begun to freeze, it has gane to a cave at B (see map), following the path with the solid dotted line. Now it first sets off, but then stops. *Constructs a plan for action:* Wait, it will be quicker to take another way. I have noticed that there is a hole in the Nettle-hedge at the big Kay-bush, so I will first go in the opposite direction. Then I will go up the hill and straight. I will go to the Plumb-Berry-tree and then takethat shortcut through the Nettle-hedge.

Example (1.3): 'Arrange a place for sleeping and pass an obstacle'

M(tired; -), PE(Is at a Shell-Berry-tree in the eastem part of the area, where it has been eating Shell-Berries, it is very windy, wind coming from south)

Usually when the creature is in this area and gets tired, it first goes either to A or to B to get a twig from a Long-Thread-bush (which it needs for arranging a nest in an Umbrella-tree). Where it goes and which path it chooses to get there depends upon the distance to A and B respectively, which ditches are waterfilled and upon how the wind is blowing. *The creature plans:* I am of course closer to B, so shall I go there then? I will run into a he-adwind; and the wind is very strong. It is extremely difficult to mave in a strong head-wind. Thus it is better to go to A; walking along the Nettle-hedge. It will be longer, but there will be almost no walking in the headwind. Wait - suddenly the creature remembers something that it noticed when playing in that area yesterday: there was no water in the ditch that runs from the Lind-bushes to the Umbrella-tree. So to takethat way would be still better. - The creature sets off. When it arrives at the Lind-bushes, however, a big Thom-tree has fallen down and lies over the path. This is something the individual has never experienced befare.

Illustration 6. An unfamiliar obstacle

M(tired; -) PE(At Lind-bushes, a big Thom-tree lies over the path- an unusual obstacle)

The creature *constructs aplan* for how to deal with this: I cannot*jump* over it - it is too high and too broad. Climb over it as over a fallen Red-Berry-tree? - no, it is too thorny. Hit at it like hitting at a Nettle-hedge?...No. Pass under it? Toere is too little space between the trunk and the ground. But maybe if I first dig a bit under the trunk...(Remember that 1-Creatures normally dig to increase the size of caves.) Yes, I will try this.

Example (/.4): 'Deal with an enemy'

M(fear; -), PE(strong smell of enemy, no tree is near)

Thinks about what to do: Toere isa High-tree behind the hill over there, but that is too far and there is no tree nearer. Where can I hide then? Behind the Kay-bush? Maybe I shall try to get a stick for attacking - if it is a small enemy - just· over there? Yes, I will do that first and then go behind the bush.

Comments: all these examples are examples of *immediate planning*. In the first example the individual is hungry but there is no food immedately available, in the second, it is cold but there is no way for it to immediately do something about this. In the third example, the 1-Creature is tired but cannot immediately deal with this, and in the fourth example finally, the individual is threatened and does not know at _once how to deal with the situation. (You may also note that these four examples to some extent are parallel to the four first examples on page 43-44, where we had the person who wanted to have something to eat with her tea, the freezing tourist who wanted to get back to his hotel to get warm, the woman who wanted to get home at night and planned to climb a wall to catch a bus, and finally the man who constructed a plan to deal with some threatening youngsters.) As the examples are instances of immediate planning, the problems that the planner deals with are urgent and thus 'problematic'. The *difficulty* of the problems, however, differs between the situations described. To plan how toget toa sheltered place or to a sleeping place from where one is, and how to deal with an approaching enemy, are *recurrent* planning problems for an 1-Creature, whereas the unfamiliar obstacle in example (1.3) introduces a more 'problematic' aspect. Also, the first example has a pronounced problem solving character. The planner thinks of how to reach a certain goal *if at all possible,* and not only of *which way* to chose to reach it.

2. Why the Perceptual-Behavioural System Cannot Plan

Crucial to the *planning of action* is the fact that the planner represents what is *not perceived* at the moment; that is, situations that are now *not perceived* and, in particular, actions that are now *not performed* and thus *not perceived.* The planner 'only thinks' of actions, their prerequisites and consequences. For instance, when *planning* how to get some of those nuts up in the tree - how to jump into the tree, pick some nuts, take them along to a particular site, get a stone of a certain size, hit the nuts with the stone, etc. - one does not actually jump into the tree and perceive this, go to the particular site and perceive this or lift the stone and perceive this, etc.

Already, the perceiving-and-behaving system that I have described *can* to some extent deal with what is not presently perceived. Toere is the anticipatory moment in perception and the activation of memory perceptions (see pp.77-79). However, these memory perceptions and anticipations are always intimately related to *ongoing behaviour,* since their role is to generate and direct current behaviour. First of all, they are used as references to be reproduced as present time perceptions, and secondly they are to some extent used to fill in expected but missing perceptions. Both of these uses concem ongoing perceiving-and-behaving.

For instance, a system controlling a reference for a program action does indeed have an activated memory perception standing for a certain structure of anticipated situations, subgoals and particular actions. But, it is *not* able to use them except in the context of matching them with current perceptions to direct subsequent behaviour, or in the context of filling in current missing perceptions. A perceptual-behavioural system does deal with *representations,* even *abstract* representations, it does *act upon goals* and *attain goals.* But it never generates representations of actions or goals that it *evaluates* and assents or dissents to before it acts. The program-govemed perceptual-behavioural system can act upon a program, that is, it can act upon an overall goal and certain subgoals that are activated one after the other, but it cannot *construct* and *test* a program. This involves dissenting or assenting to proposed actions and subgoals. When acting upon a program, the system moves through the program structure, comparing anticipated - sometimes alternative - perceptions with present time perceptions. Particularly important are perceptions that represent anticipated *preconditions for actions,* like '*if the Nettle-hedge area is limited, then* hit at Nettle-branches, and anticipated *outcomes of actions,* like' *hit at Nettle-branches until* they bend'. It is on the basis of such comparisons that the system moves on, entering particular branches of the program structure. Yet, in this context, the steps are first action, then evaluation of that action, whereas for planning, there must be first an evaluation of a potential action, then - possibly- action.[26] The planner must be able to operate with representations that relate to actions that are *not* ongoing, and possibly are not going to be executed at all. It must be able to suggest an action and test this represented action

[26] This does not entall that there cannot be an alternating between planning, acting and re-planning.

as to its predicted consequences in relation to some goal. (To *think* about, for instance, whether to takethat ball or not, and about posibilities of *how* to do it.) [27]

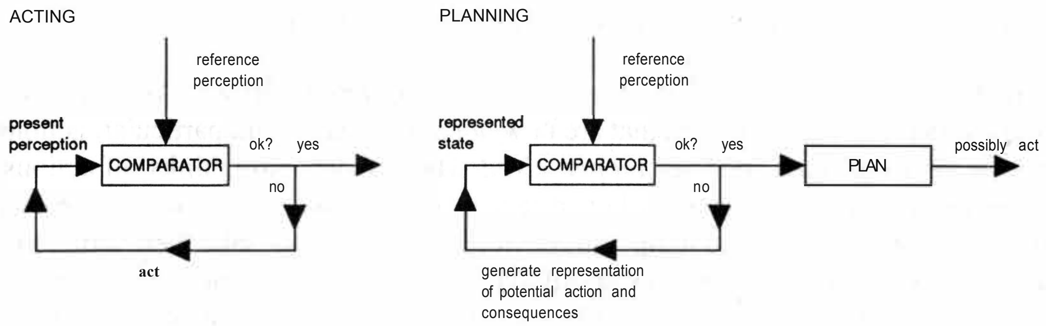


Figure 16. The role of reference perceptions in action and planning respectivefy

We want a system that can generate and test (structures of) action references or instructions without any attempts to immediately act upon them. We want an *internal* evalution and possible revision of references for actions before executing them and not just automade and immediate acting upon an activated representation. The basic perceptual-behavioural system, however, is not capable of this.[28] It does act upon various goal representations, but these are never *questioned* or *tested* in relation to one another.

3. lmagination

How can we construct a system that operates on reference perceptions - action references - that are *not* so intimately related to *ongoing behaviour andperception?* In order to solve this problem, I will consider what Powers speaks of as *imagination* or 'the *imagination mode'*. The basic design of a perceptual-behavioural system, revised to include a capacity for imagination, is presented in figure 17 (adapted from Powers, p.221.):

27on the other hand, some of the representations that the perceptual-behavioural system *does* use and act upon can, and ought to, be left out when planning. A representation of the*particu/ar* place where you tind baking powder in a particular shop, and the exact arm-movement for stretching out for and taking such a package need not be involved in your plan for baking a cake and going to the shop to buy ingredients.

28The perceptual-behavioural system does not pass Anderson's 'litmus test of planning' mentioned on page 53. It cannot first represent a sequence of intended actions and then *note a conflict* in the sequence and therefore *reorganize* the sequence in anticipation of this conflict.

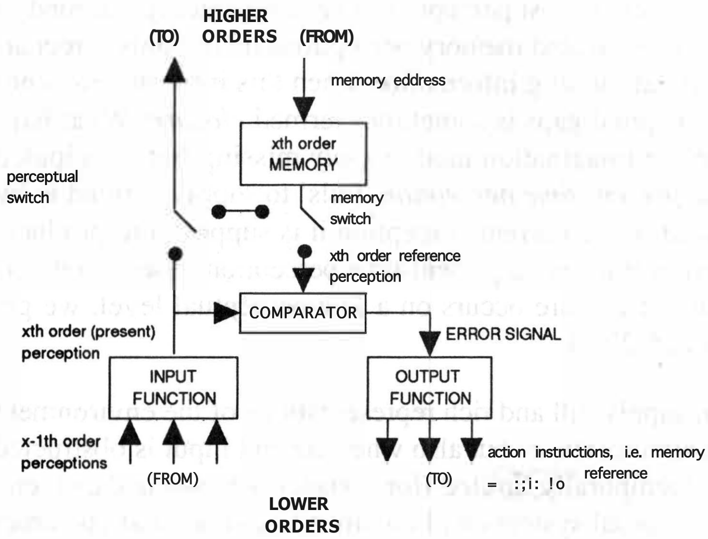


Figure 17. Perceptual-behavioural control system with imagination ability

When the system *imagines* something, *stored perceptions* at some perceptual level (more precisely, in some perceptual systems on that level) are retrieved from memory and then passed on to higher levels, just as if they had been 'normal' current perceptions coming from lower levels. The *memory switch* and the *perceptual switch* in these systems are both *horizontal* (see figure). They operate in the *imagination mode:* they retrieve memory perceptions but do *not* engage in reproducing these perceptions by action, i.e., by instruering control systems on lower orders. For instance, we can consider some perceptual systems that are in the imagination mode and activate the following memory perceptions: 'arm approaching round object', 'arm moving forwards', 'joint bending' and 'hand moving down'. The systems do not send these configuration and movement perceptions further down the hierarchy (and thus *act upon* them), but these reference perceptions are only sent on to higher-order perceptual systems that deal with them in the usual way and produce higher-order perceptions, for instance, the routine perception 'grasping an object'. Or, as another example, the memory perceptions 'stick hitting nettles' and 'jumping over something' may be *imaginatively* produced andresult in higher-order perceptions like (parts of) a getting-over-nettles-program or a jump-over-obstacle-program. It is of course also possible to retrieve lower-order perceptions out of which the higher-order systems will not make any sense - that is, out of which they will not produce any higher-order perceptions or invariants.

As far as levels above the imagination switch are concerned, imagining is *the same* as perceiving or as perceiving-and-behaving. A past perceptual situation is recreated. But the experience as a whole differs: as the levels below the imagination switch are not involved, there is a lack of details and 'thickness' of experience. For instance, in the second example above there are no representations of separate movements, tree texture, colour, form of the nettles, colour of the stick, hardness of the tree, muscle efforts, pressure in grasping, etc.[29]

[29] The character of the experience depends of course on ***the leve/*** in the hierarchy that is involved and the ***number of sub-systems*** on this level that are in the imagination mode.

Toere is an *aspect of imagination* in all or most perception. I have repeatedly returned to the *anticipatory* factor of perception. Activated memory perceptions do not only direct attention but may also *add expected* but missing information when this information is not directly supplied. Pilling in of perceptual gaps is sometimes termed *closure.* What happens in closure is that systems in the imagination mode *supply* missing but anticipated perceptions. More precisely, what *present-time perception* 'fails' to supply is filled in by a reference perception itself instead of the current perception it is supposed to produce. The reference perception is treated *as if it* were a present-time perception based upon perceptual signals from lower levels. (If closure occurs on a *low* perceptual level, we get *hallucinations.)* (See Powers, pp.225-226.)

Thanks to closure, perception can supply full and rich representations of the environment not only under ideal or favorable circumstances but also when current input is obstructed or scarce, when it is spatially and temporally limited (for instance when it is dark), etc. Under such circumstances the perceptual systems will corttinue to extract and construct important perceptual invariants. (Cf. the discussion, page 76-77, of certain perceptual anticipations as the internalization of reliable constraints in the environment - internalized through learning or incorporated into the genetic code during the evolution of a species.)

Imagination and closure are really the same kind of phenomena. In closure, however, we deal with comparatively short periods of time, and thus there is a limited determination from within. Present-time perceptions that are 'unexpected' since they constitute a divergence from the perceptual picture shaped by anticipations and closure - the 'closed perceptual picture' - will, if they insist, make their way through. This is not what happens in imagination. In imagination there can be long successions of memory perceptions which evolve according to their own rules and which are not compared or matched with *any* current perceptions. For instance, when replaying or activating a program perception in the imagination mode, no comparisons between any present-time perceptions and the structures represented in the program occur. Toere is just a long sequence of *unbroken* closure - information supplied entirely 'from within'.

Furthermore, in the imaginative mode, reference perceptions are more easily combined, transformed, replaced, etc., than in the perceptual-behavioural mode. An 1-Creature may, for example, at one moment imagine 'eating Red-Berries at the edge of the woods' and the very next moment imagine 'jumping over the river', whereas it is not possible to entertain these 'references' as present time perceptions in close succession.

The *main difference* between the two modes is that in the *perceptual-behavioural* mode a replayed reference or memory *perception forces* action to occur: the retrieval or activation of a reference signal or memory perception instantly initiates action in order to reproduce the memory perception also as a present-time perception. (Powers, p.219.) In imagination the activated memory perceptions have no function as *reference* perceptions. They are processed just as if they were present time perceptual input from lower orders. Furthermore, *behaving* requires that control systems from the highest down to the *lowest control level* actively operate to control perceptual dimensions and reproduce perceptual references. In *imagining,* on the other hand, only the systems above the memory and perceptual switches are operating. Imagining some action or event does not have to involve representation of all details. (Powers, pp.217-224.)

The imagination mode in this model is clearly *distinct* from the perceptual-behavioural mode. During a certain time, the system imagines *instead of* - or at least *at the expense of* - perceiving and acting. A particular control system is *either* in the imagination *or* in the perceptual-behavioural mode. (Powers, p.219.) In this context I want to emphasize again (cf. p.39) that internal activities, such as imagination and planning, are not 'for free'. They are resource-demanding and compete with other activities. It is not possible to engage in imagination, thinking, planning, and so on, and simultaneously to fully engage in perceiving-and-behaving, that is, to perform advanced perceptual analyses and produce advanced behaviour.

I will expand on this as it is an important point for my thesis. An individual that is engaged in imagining cannot perform just any actions. Only *certain kinds* of behaviour can occur while imagining. If the memory and perceptual switches are set horisontally on the program-level - that is, if imagination occurs on the program level - related *routine actions* may yet be performed and *automatized* behaviour can occur. (It may be possible, for instance, fora creature to walk while imagining jumping up into a tree.) The reason for this is the following: That a behaviour is automatized means that the *perceptual switch* is *open* (not vertical) and that perceptions of this behaviour are not passed on upwards. No *higher-order perceptions* of the behaviour occur, and therefore the higher-order systems in question are free to do something else: deal with the consequences of the behaviour, deal with other aspects of the perceptual-behavioural situation, or engage in imagination. (Cf. Powers, p.222.) In general, behaviour organized on levels *be/ow* an imagination switch *can* take place *white* imagining. Furthermore, there are *many systems* on each *leve/,* and *some* of these may be in the perceptual-behavioural mode and *some* in the imagination mode. Thus it is possible, while imagining *certain* things, to perceive and do *certain* things but *not others.* For instance, several experiments on so called interference tasks show that a visual perceptual task is more difficult to perform than an auditory perceptual task while visually imagining something, and that the auditory task is more difficult than the visual while auditorily imagining something.[30] (See for instance Neisser, 1976, pp.145-149.) In sum, imaginative actitivy competes for resources with action and perception. Only *certain* kinds of perception and behaviour can occur simultaneously with imagination.

Experiments as those just mentioned above - and many others - indicate common resources for *imagination* and *perception.* And indeed, that there *are* non-trivial relationships between imagination and perception is hardly questioned. But there is a great deal of controversy as to the specific nature of the relation. According to the current model, it is basically *the same systems* that are involved in perception and imagination and partly the same mechanisms. In particular, on the levels above the imagination connection, the perceptual signals are identical. Yet the phenomena and experiences as wholes are *distinct* from each other. (Cf. p.97.) And normally we *know whether* we are perceiving something or imagining it, even though there are 'unclear cases', like in hallucinating.

In imagination it is thus possible to *reproduce* a *past* perceptual situation *without acting* to set oneself in the corresponding 'extemal situation'. Hence, it is possible to 'evoke' other

[30]And recognition and discrimination in perceptual tasks is influenced - positively or negatively - by prior imagination.

perceptual environments than the actual one. A Berry-Creature sitting on the ground can, for instance, imaginatively reproduce a situation where it jumps into a tree or when it holds a berry in its hand, without moving a limb. Or it can reproduce the route from place A to place B without actually traveling it. Nothing changes in the system's *external environment,* but there is *simulation of change:* simulation of certain events and actions. The changes take place in an *internal model of the external environment* - or in an *internal environment.* Reference perceptions of different kinds can be said to model aspects of the external environment if they relate to each other in ways that correspond to spatial and causal relationships between elements in the external environment.[31] A program reference, for instance, may model certain rules in the world by the represented outcome and precondition relations of the actions as well as by the sequence or pattern of those represented actions. A routine reference such as 'moving from stone site x to hill y' models a certain spatial structure by being invariant over a number of memory perceptions of particular paths. Particularly important are the memory perceptions that link certain objects and activities to particular sites and those that relate such sites to each other. A *spatial representation* of the external environment, that is, a representation of *where* certain things are, and how these sites or places of interest relate to each other, is a fundamental aspect of an internal environment. In ethology, as well as in psychology - animal and human - the notion of a cognitive map is a frequently used notion. I will say more on this in next chapter.

Because of the dynamic aspects of the internal environment - the activation of elements and the evolution of functional states according to particular rules - it is proper to speak of *simulation* of processes or changes. Internal changes occur, and relationships between internal elements develop in ways that, to some extent, correspond to events and actions in the environment and changes in relationships between external elements.

Imagining can be seen as a kind of more or less random *exploration in one's internal environment* or of one's knowledge. It is not a search in order to direct ongoing behaviour but a nondirected exploration for no (immediate) behaviour control purposes.

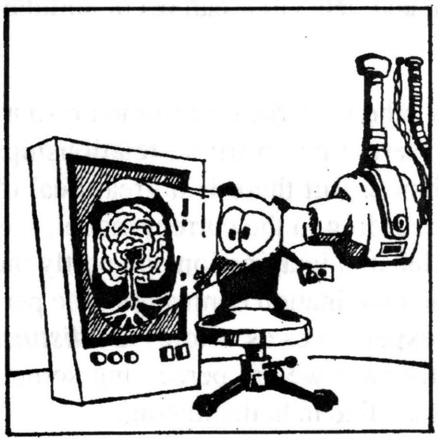


Illustration 7. Internal exploration

[31] Some activated memory perceptions will evolve into other memory perceptions according to the same rules that make the world evolve from one moment to the next.

The *creative* aspect of imagination is important. Even though imagination is, *on some level,* an issue of *reproduction* of past perceptual situations or experiences, on higher levels resemblance to past experiences can be lost. Memory perceptions can be *reorganized* (elements combined, reordered, deleted, inserted, replaced, etc.) and *novel combinations* of memory perceptions can be produced. (Cf. Powers, pp.223-224.) A Berry-Creature may for instance imagine a red Shell-Berry although it has never perceived any, or imagine a tree at a particular site where it has never perceived a tree. (And as noted before, p.97, it may also happen that higher-level perceptual functions cannot make sense of some set of activated memory perceptions from lower-order memory systems.)

4. Planning as Internal Generating-and-Testing

It is obvious that the considerations in the previous section provide some basis to build upon when we want a system capable of making plans for action. I have introduced a capacity for imagination - for generating perceptions that are not related to ongoing behaviour, and for perfonning operations in or 'exploring' an internal environment. In particular, the capacity involves an ability to internally_*reorganize* and make new combinations of memory perceptions.

But it is also evident that, in order to have a planner, more is required thanjust adding an imagination switch to the basic design. Planning is not just an issue of imagining situations and actions in a random manner. It is not an aimless succession of activated memory perceptions. We want a more selective and *goal-directed* imagination with some evaluation or testing of imagined events and actions in relation to certain represented goals.32

Planning one's action, I suggest, is in principle a question of *internal or imaginative generating-and-testing* of action instructions for oneself. We have already, in the previous chapter, pp.87-88, considered *external* generating-and-testing of actions, namely, trial-and-error behaviour. In this context we have the following components:

- A *start state* given by the system's motivational state plus its perceptual environment

- Some *end state(s)* or end condition(s); i.e., states where the trial-and-error behaviour is terminated (which consist of a state where the motivation ceases ora state where the system can apply a routine ora pre-formed program)

- The system's *repertoire ofbehaviour operations*

32A requirement that is not essential for imagination as such but is crucial in planning is that the system's intemal environment is ***rich and correct enough*** for representing the outer environment. A stone for instance is rigid, heavy, hard, nonpenetrable, and consequently also in the intemal environment some transformations of stones must be allowed and others forbidden. (Stones do not for instance melt as wax if you warm them up.) Objects and events must be represented with (1) adequate properties and affordances, to use a Gibsonian term, and (2) with those properties and affordances that are of interest to the system. (On the other hand, these are also requirements for an imaginative capability used in ***anticipating*** and in ***closure.)***

A *trial* consists of a behaviour operation, i.e., a movement, a routine execution, an execution of part of a program or an execution of a whole program. Reaching an end state means success, not reaching one means *error*.

If we now turn to planning, the following components are required:

• A *start representation,* i.e., a representation of the initial situation from which action is planned to take place and a representation of the problem, i.e., the relevant discrepance between the start situation and the goal situation

• A *goal representation*

• A *repertoire of representations* of (1) *situations,* more or less schematical, and of (2) *processes* (transformations between situations), in particular of *actions* - with their preconditions and consequences - again more or less schematical

All these elements might be *memory perceptions,* i.e., stored perceptions of situations and processes (actions in particular).

Metaphorically one may think about planning as making a film. The repertoire of representations is a container filled with film sections of varying lengths and details that represent situations or processes (events, actions) - like 'walking', 'climbing over something', 'holding a nut', 'holding a stick', 'cracking a nut', 'being on top of a hill', 'hitting something with a stick', 'walking to town', 'catching a bus', 'walking into a room', 'looking under a sofa', and so on. Planning starts out from film segements that correspond to start and goal representations. And between those, the film, and the plan, are to be constructed. Film sections and sequences of film sections are retrieved from the box, run, and 'tested'.

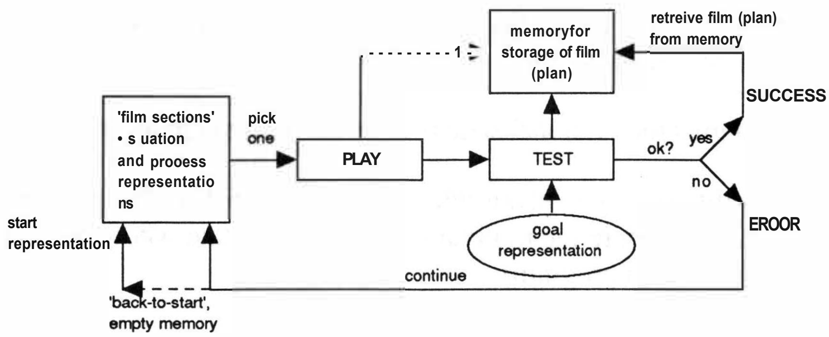


Figure 18. Planning as the composing of a film

Osing the film composition metaphor, two important differences between extemal - behavioural- and intemal generating-and-testing can be illustrated. One concems the possibility of 'starting afresh', of 'regretting and beginning anew', i.e., retuming back to start. In intemal trial-and-error- but not so in behavioural - it is *always* possible to return back to a relevantly similar or even identical state - and to do so instantly. We can even have a special operation 'back-to-start' accomplishing this. (See figure 18.) In the metaphor given, this corresponds to getting rid of the films we have composed so far. This may, for instance, be done after a certain time, when a trial has not been successful. No intemal trials can have the consequence that there 'is now no way back'. *This* is, however, something that can happen in trial-and-error behaviour.

The other, and related difference, concems the fact that in intemal generating-and-testing it is possible to operate with several parallel trials. It is quite feasible to work on more than one sequence in parallel - whereas in behavioural trial-and-error, only one sequence can be run at a time. *Internally* we may have *many* generate-and-test processes going on simultaneously.

However, there are some important flaws in the film metaphor as I have presented it. It describes the repertoire as justone big box, where *all* the system's memory perceptions are contained. But, such a generate-and-test system cannot be very effective. There is likely to be a *massive amount* of *search* for a successful trial. Because it is evident that once the repertoire of operations is rich enough, only a few trials can be appropriate and most will be errors. The 'film making' will be a very resource consuming enterprise.

In extemal trial-and-error it is not *physically* possible to apply an arbitrary behaviour operation in a given situation. But in imagining, that is, in search or exploration in an intemal environment, one can, in principle, retrieve *any* piece of knowledge and imagine or think of *anything.* It is true that an inherent advantage of intemal trial-and-error compared to behavioural trial-and-error- or of simulation of operations compared to actually carrying them out - lies in the possibility of *speeding up* operations.[33] But with a repertoire that is rich enough - and it will be so in any interesting and biologically realistic system - this advantage is of no help. And of course there is no point in having a massive or even never ending inner search replacing behavioural search, even if the latter may take time as well - and be dangerous. On the other hand, there *has to be some kind of searching:* the system does not know bow to act and is looking for a plan for action. I will suggest some ways as to bow the memory search and the generating-and-testing of plan elements may be restricted to limit the search space.

(1) Planning starts from *what is familiar* in the planning situation.

A planning situation is, as I have mentioned a number of times, a situation that is *to some extent unfamiliar,* and, therefore, an individual does not know immediately what do to or bow to act and does not immediately start acting.[34] But, on the other hand, a planning situation must in some respects be *familiar.* Because if not, a planner has no chance at all in dealing with it. Thus, more or less detailed features of the situation are perceived and recognized by the planner, and *this,* in the first place, we let direct the search for retrieving memory perceptions and in particular memory references for action or behaviour elements. *Memory* is indeed far from being 'just a big box', but is distributed sothat each control system has a memory (sub-)system of its own. The fact that the organization of memory is *developed* for the purpose of *perceiving-and-behaving* ensures that there are certain *associative links* between memory perceptions, within as well as between memory

33some such speed-up facilities originally used in perceptual anticipation will be available.

34It is conceivable that the pianoer constructs a plan that does indeed *concord* precisely with a preformed program. (And thus its behaviour will look 'as usual'.) Yet insofar as there is generating-and-testing to arrive at this, it isa case of planning. The plan/program is *constructed now,* to fit this 'novel' situation.

systems. This may now direct activation of memory perceptions, so that, for instance, a system planning how to get some berries will retrieve *plan elements* from memory systems in control systems *normally dealing* with food gathering, rather than generate plan elements like the sensation of water, the digging of a hole, the blue sky; etc. As a particular control system handles particular *kinds* of actions, events, routines, configurations, sensations, etc., search in the memory of this system ensures activation of 'related elements', and there will further be associative links to other systems on different levels.

Yet, *new* 1.deas, unusual elements of a plan, *can* enter into the planning process; first, from long chains of association and second also possibly from some kind of wilder activation or retrieval of memory perceptions. And, third, they can, in *particular,* be obtained from *novel combinations* of 'usual or habitual' elements.

To continue with the idea of using what is familiar; at least in some planning domains the actions of a planner will vary much less than the situations where its plans are applied. Here planning may basically be a question of retrieving a *preformed concrete program* and modifying or *adapting* this to meet the novel situation. (Cf. Alterman, 1988.) Consider the example with the Red-Berries growing very high up. Here the individual may start out from its preformed Red-Berry-program. Simulating this, it may notice the divergence and failing precondition as to the reaching of the berries with a usual stick. And so it may interrupt the program structure at this point. In the situation, as well as in the program, there are several elements to initiate associative retrieval of plan elements from memory for replacing the unsuccessful part of the old program or plan; for instance, associating from stick or digging to the routine for wrenching loose sticks for manipulating Umbrella-trees. And in this way new sequences of behaviour elements may be represented and tested - given that there are also methods for replacing, adding and deleting program elements. From the Red-Berry-program it will also be possible to reach the more *general* Berry-program and from this to go to other *instantiations* of this general program. Two processes that indeed seem essential in the constructing of a plan for action are the *generalisation* and the *specialisation* of representations of behaviour elements. Generalisation or abstraction involves removing details from such a representation, a program say, in order to have more general or schematical invariants, (like generalising the Red-Berry-program to the Berry-program). Specialisation involves tilling in details, that is, creating more concrete and specific instantiations of a general invariant, (like specialising ' passing over' to 'climbing' or 'climbing' as moving ones limbs in a particular way or as moving to particular sites).

(2) Planning utilizes the hierarchical and distributed organization of perception and memory.

The processes of generalisation and specialisation are grounded in the perceptual hierarchical apparatus. Furthermore, just as *in perception* many perceptual systems on many levels *simultaneously* do their work, *plan construction* - suggestion of elements, parts of the plan, etc. - can occur *simultaneously in several systems* where some deal with particular kinds of locomotion, some with stick manipulation, some with the visual aspects of a kind of area, etc. They can also occur *at several levels of abstraction;* for instance, when a planner simultaneously thinks about whether to flee or hide and whether to get a tool *and* thinks about details of tool use - or simultaneously thinks about whether to go to a Kay-bush or to a Lind-bush *and* thinks about where there is a Lind-bush.

I do not think that so called *hierarchical planning* models capture planning processes in biological planners particularly well. In such models, plan construction is a purely *top-down* process. One starts by building the abstract plan and then, by a process of specialisation or refinement, systematically moves down the hierarchy to produce more and more detailed and concrete plans. *Bottom-up* processing, on the other hand, means to elaborate a part of a plan or a sub-plan, while still lacking the overall plan. It would be like setting off to plan bow a Kay-branch might be used to retrieve berries before planning bow to *get* a Kay-branch (and *maybe* realizing that the nearest Kay-bush is really far away.) This kind of processing means starting out from *apart* of a possible solution and to extend one's plan from there, possibly to another *part* of a solution or plan. (Cf. *The Handbook of AI,* p.519.)

My doubt as to the adequacy of hierarchical planning models can be related to Hayes-Roth's and Hayes-Roth's research on what they call *opportunistic* planning (Hayes-Roth and Hayes-Roth, 1978). Opportunistic planning involves observing an opportunity for dealing with one goal or subgoal and that the planner goes on to construct a sub-plan for dealing with this but possibly without regard for how this sub-plan will integrate into an overall plan. This, Hayes-Roth and Hayes-Roth maintain, often happens with human planners. They also model the ability of human planners to reason at many levels of abstraction and to move freely between them" in both directions.

Planning in biological systems, I believe, is quite probably both 'top-down' and 'bottom-up'. Which mode dominates will depend on the actual planning problem and planning situation.

I the context of bottom-down-processing, two things are important. On the one hand, there can be *generating and testing simultaneously* in many systems and on many levels of abstraction. On the other hand, such parallel processes can influence each other to some extent, so that if there is a 'good idea' in one system and on one level this may be signalled to other systems and other levels. In testing, assent may be required from more than one system and level to a particular element suggested - otherwise it is discarded. Furthermore, if there is much processing done in parallel and elements are tested by more than one tester and not so easily allowed to develop, this means that we can afford suggestions of 'real stupid elements', the irrelevant and inadequate suggestions. In other words, we can afford to let *error* in as well as the crazy suggestions that may turn out to be ingenious.

Now, as the idea is to obtain planning by introducing imaginative generating-and-testing in a basic perceiver-and-behaver, a question may arise as to whether the perceiver-and-behaver in question must be of a certain *order.* I will try to answer this briefly. Consider a routine-govemed system. Here, remember, genetic information and/or the systems perceptual environment moment-to-moment determines, at any given moment, the routine action - some end condition and an overall pattern of actions - that is to be performed. It is not difficult to see that this leaves very *little* room for *planning.* The system cannot represent *alternative* routine actions, test them and chose to act upon one of them. All that might occur is some planning within the frames of a routine, that is, of particular movements and details. But the activation of a routine reference implies that the system acts upon it, and behaviour units below the level of routine actions are often swift. At those levels of behavioural organization there may simply not be much time for planning.

(There is, remember, also not much space for *behavioural* generating-and-testing on lower levels of control, as the goals and problems on these levels change swiftly and that there is thus no relatively constant situation in which to test *alternative* actions or movements.) (Cf. pp.87-88.)

Furthermore, and more important, it may be doubted that planning - intemal proposals and evaluations of actions and goals - at such detailed levels is a good way to produce action. Instead of planning exactly bow one's armshall move tograsp an apple, i.e, to represent various possibilities for this whole motion and decide on one, it is hetter to start moving the hand. Toen, action references on various levels will be immediatley *acted* upon, and there can be a continous adjustement in relation to the local environment. In this way the structure of the environment and one' s operations upon this may be a 'stand-in' for highly detailed memory stores and costly information-processing operations. Therefore, my conclusion is that planning is not interesting - it is not useful and not required - until the level of *program-governed* systems.

Thus, the general framework for modelling the planning of action that I propose is the following. Toere is *simulation* of actions and events by osing the perceptual apparatus in the *imagination mode.* By means of search in an *associative memory* and *generating-and-testing* of (sequences of) retrieved memory perceptions - on several levels of abstraction and in several control systems - plans or plan elements are suggested and tested. Planning primarily occurs in program-govemed perceptual-behavioural systems. Finally, the planning of action, as a kind of imagination, is performed *at the expense of action and perception.*

5. Arguments For and Against a Generate-and-Test Mode/ of Planning

The idea of planning as a kind of intemal generating-and-testing is not particularly original, and there are good reasons for this. First, it seems a plausible model when thinking informally about planning. To plan one's actions means to make use of knowledge and experience from past courses of actions in order to structure future ones. (Cf. Dennett, 1984a, in his discussion of intelligence.) From the knowledge that one already has, one *generates* various plans and parts of plans and *tests* those against certain criteria.

Second, if one moves to formal models, like AI models of various mental activities and competences, one will tind generating-and-testing to be a ubiquitous strategy. In a hierarchical manner, the competence or capacity to be explained is broken down into a generator anda tester, where the former generates candidates for solutions, or elements of solutions, to the latter, which, in turn, accepts or rejects these candiates on the basis of some criteria. And this process continues until it arrives at task descriptions that are obviously mechanistic.

But let us consider two kinds of criticism of the use of generating-and-testing for modelling human mental activities. The first contends that generate-and-test-models are *too mechanical and simple.* A reply to this, however, is that generate-and-test-models *need* not be either simple or *obviously* mechanistic. lndeed a generate-and-test mechanism will not be efficient or powerful unless the generator is endowed with a high degree of selectivity. This can be obtained by having *several layers* of generating-and-testing. (See Dennett, 1978, p.86; cf. also the discussion above of the use of the hierarchical organization of perception and memory in planning.) The second type of criticism is that generate-and-test-models *idealize* human beings: people do not follow some ideal Popperian procedure. They do not try to produce various alternatives and to falsify them but are much more conservative. They strive at finding and verifying *one* alternative (often a favoured idea or solution). The replies to this challenge in the context of planning are the following:

(1) There are different kinds of planning (and problem solving in general). Sometimes one may be after just any solution; just *a plan* for action that will do. At other times one is *considering alternatives* and looking for *the best plan* for action. But quite apart from this, a generate-and-test-model of planning does not have to be a model where what is generated and tested are *alternative entire plans.* It can be elements or parts of a plan. Toere can be generating-and-testing of subgoals, of representations of the problem, of possible constraints of the solution, etc.

(2) The model is not confined to processes and events on those cognitive levels that are accessible to introspection or consciousness. It is likely that there is much generating, testing and rejecting of cognitive elements that is *not* accessible to introspection.[35] (Indeed *all conscious* thoughts and ideas (that perhaps are consciously tested) might be *results* from unconscious generating-and-testing procedures.)

Now assume that the viability of generate-and-test models of psychological competences is in good standing. Still you may think, though, that it is just one kind of model amongst *other kinds* that are just as sound and viable. In "Why The Law of Effect Will Not Go Away", (Dennett, 1978, pp. 71-89.), however, Dennett argues - to my mind convincingly - that this is not the case. I want to briefly dwell upon this account, which is an account that not only describes planning as generating-and-testing but situates it in a *broader framework* of generate-and-test processes. My main motive for doing so is that I believe that the virtue of generate-and-test models as an account of planning - and other mental acitvities - thereby becomes clearer.

By analyzing the similarities and possible causal relationships between *natura/ selection* and the *law-of-effect,* Dennett distinguishes *four levels* of the principle of generating-and-testing in behaviour control. See figure 19.

35Consider again the objection that generate-and-test-models are too mechanical and stupid, whereas *people* are not so stupid: people do not generally come up with *large quantities* of *irrelevant and stupid* suggestions and ideas. This can only be used as an argument against generate-and-test models if one assumes that the generating-and-testing takes place solely on a conscious level. I do not assume that and, as I will shortly return to, I see it as the main virtue of the model that it *admits stupidity.*

testing mechanism	*elements that are generated and tested*
• natural selection	• genetically determined behaviour-producing mechanisms
• learning in an individual (learning by experience)	• not genetically determined behaviour and behaviour-producing mechanisms
• planning, thinking, in an individual (learning by thinking)	• potential behaviour and potentially behaviour-producing mechanisms
• planning, thinking, in an individual (learning to learn)	• planning, thinking and learning mechanisms

Figure 19. Four levels of generating-and-testing

On the first level, what is *generated* are individuals or species with different kinds of genetically fixed mechanisms for producing behaviour in particular circumstances: instincts, tropisms, taxis, etc. Natural selection *'tests'* these in the sense that (the most) adaptive mechanisms are proliferated, others are not.

In the next step some mutations appear, and we get creatures with *reinforcers* - where a reinforcer is some physical event that increases the likelihood of the recurrence of events that produce them. Some mutations may have reinforcers that are events normally produced by biologically useful situations and events like the presence of food or water, increasing distance to an enemy, etc. Others may have reinforcers that are events normally related to injury and deprivation, etc. Now, natural selection again *tests* these different mutations that have been generated, in the sense that those with adaptive reinforcers will proliferate. With this - and with a capacity for *trial-and-error* behaviour - it then becomes possible to have *generating-and-testing* of kinds of behaviour and forms for behaviour control *in an individual system.* Through the reinforcing-mechanism, certain kinds of behaviour and behaviour control will be *selected* or *learned.* The likelihood of the recurrence of a *certain kind of behaviour* [36] in a certain type of situation is increased.

It is in the *third* step that we get *internal generating-and-testing* - that is, creatures capable of generating and testing types of behaviour not by performing them but by 'thinking about them'. Now, individuals may *test potential* actions by simulating them in an *internal environment.* They can produce and select elements of *potential* behaviour control - like plans, preferences and expectations - which, in turn, will make certain kinds of behaviour likelier in certain kinds of situations. Dennett's account of the evolution of *planners* (or in general of thinkers) - that is, of the emergence of this third level of generating-and-testing - goes like this. We get *mutations* endowed with *inner environments;* that is, inner regions in the brain where physical events can affect and are affected by elements of *potential* behaviour controls (like beliefs, expectations and plans) and through this certain

36 And it is here that natural selection works - in the end, it is *always behaviour* that is 'tested' and selected (cf. page 19).

forms of behaviour are made likelier in certain situations.[37,38] Mutations with inner environments that reinforce non-adaptive behaviour - by selecting false expectations, unsound plans, etc. - will not survive, but those with adpative inner environments will. An inner environment can be maladaptive by using inadeqate test criteria, incorporating false expectations, etc. *(Ibid.,* pp.79-80.)

Dennett conceives of the inner environment at the third level as *fixed.* In a fourth step, however, he introduces mutations leading to creatures where the inner environment as such may evolve in an individual. And this occurs, again, by means of internal generating-and-testing, now of elements of the internal environment. (We have not only learning by behavioural trial-and-error and learning by planning and thinking, but we also get a possibility to *learn how to learn* and *learn how to plan and think.) (Ibid.,* pp.78-79.)

What are the virtues of this account? And what is *special* about generating-and-testing as a model of cognitive and behavioural capacities and of their evolution and development? Its main virtue, I claim, with Dennett, is that it gives a *non-question-begging* account of adaptiveness - of adaptive *choice* between possible solutions to various problems. It is a *non-intelligent mechanism* for generating intelligence and goal-directedness. First, *natural selection* explains the adaptiveness, the 'intelligence', of genetic constitution and biological hardware (including genetically determined learning mechanisms) *without* presupposing an intelligent *designer* that already knew or appreciated that these would be good ways of doing it. It does not presuppose that the intelligence of the solutions originates from intelligent thoughts in an intelligent creator. Second, the *'law of effect'* and the phenomenon of learning explains adaptation and development within an individual without presupposing an intelligent teacher who aldready knows the good solutions. And, finally, *internal generating-and-testing* plays a similar role when we want to account for the 'intelligence' and goal-directedness of a plan for action. We want to explain the construction of a sound, 'intelligent' plan for solving a behavioural problem (to deal with a situation that is somewhat unfamiliar) amongst *possible* plans and random collections of action representations. We want to explain this without making an appeal to an intelligence inside the system, a little action-plan-expert, that already *knows* the good answers, or can figure them out. The 'stupid' elements of thought must also be produced somewhere. Stupid things must also be suggested, whether consciously or non-consciously.[39] (This consideration counters the criticism above of generating-and-testing as mechanical and stupid.)

37If *to learn* means to increase the likelihood of a particular type of behaviour in a particular type of situation (where the change is a change in environmentally appropriate direction and is based upon some specific experiences in the organism), then planning is a kind of learning, or at least an element of learning. One learns by thinking. You do not search for solutions in your external environment but you search in your *inner environment.* And so you learn without *acting.*

38Dennett's notion of ioner environment is not equivalent to my notion of internal environment (p.70) Whereas my notion includes the total set of memory perceptions of a system, Dennett seems to be after a more selective set of representations - those that represent the criteria, against which proposed cognitive elements such as plan elements, expectations, etc., are tested. (Dennett, 1978, pp.77-80.)

39There *might* exist some individual differences in this domain regarding *at what levels* the mass of stupid suggestions appear. If you belong to the group of people that for every sound idea you get and maybe pursue further, you get a thousand (conscious) ideas that go to the trash, you may find the idea of generating-and-testing more plausible than if that is not how you experience your - conscious - thinking.

6. Conclusion

In this chapter I have sketched how the basic design may be extended into a system capable of planning. Planning is, basically, a question of imaginative generating-and-testing; an *operating* on action references or representations without *acting* upon them. The generating-and-testing may occur at several levels, using the hierarchical memory-structure. I also suggested how this hierarchical memory may be used to delimit and direct the thinking or imagination. A program-governed system may thus construct and evaluate new programs and subprograms, that is, it may plan. Finally I spoke of generate-and-test models in general for modelling psychological phenomena, and planning in particular.

It is important to see how the planning capacity relatively directly builds upon properties of the basic perceiving-and-behaving system: for instance, the capacity to construct and handle perceptual invariants, the capacity to deal with more or less abstract representations, the hierarchical structure of behaviour and perception, the goal-directedness of behaviour, the subjective and constructive side of perception, and so on. (Without these properties we would, to my mind, be left with more of a mystery around this astonishing capacity of thinking about and structuring one's own actions beforehand. An advantage of constructivistic over non-constructivistic theories of perception, is precisely that they help fill in this gap.)

In the following two chapters I present and discuss some more details of the immediate and the anticipatory planner respectively; sketching a systems understanding first of an immediate and then of an anticipatory planner.

8. AN IMMEDIATE PLANNER

1. Introduction

Following the ideas that I outlined on page 20 in the chapter 'Biological Functionalism as a Framework', I will in this chapter be concemed with what I call a *systems understanding* of the planning of action in an immediate planner, exemplified by an 1-Creature. I want to discuss the following questions: (1) *When* does planning occur? (2) How is an individual *motivated* to engage in planning? That is, what are the mechanisms for regulating - for starting and stopping - planning activity? and (3) Given motivation or overall regulation, how can the planning of action proceed? In particular, what kinds of *knowledge* must be involved?[40]

2. When Does Immediate Planning Occur? - When does an 1-Creature make Plansfor Action?

What are the *kinds of situations* in which an 1-Creature starts to make a plan for action? *When* - i.e., under what kinds of circumstances - does it engage in intemal generating and testing of possible courses of actions in order to arrive at a representation of a set of behaviour instructions for moving from a start- to a goal situation? When does it enter a 'planning mode'?

Planning situations for I-Creatures are characterized by two properties. First, the individual has *a primary motivation* and a corresponding problem. That is, it is motivated to engage in search for food and eating, sleeping, getting warm or evading a threat. Second, there is something that is *novel* - not completely familiar or foreseen - in the situation. It is *not 'just as usua/',* which means that the individual can have no routine or preformed and specified program that fits the situation, which it can retrieve and act upon.

There are of course many different ways in which a situation can be *'novel'*. In the example with the Red-Berries (p.97), the berries grow *unusually* high up. In the example where an I-Creature plans to go to a place to sleep, there is an unfamiliar obstacle (the fallen-down Thom-tree), whereas in the example with an I-Creature on its way to a windshelter, a familiar obstacle (water in a ditch) has been *removed.* Furthermore the 'unfamiliarity' or variance can be an aspect of the system's perceptual environment, i.e., something that it currently perceieves, but it can also be an aspect of its intemal environment, like when the novel element is an aspect of an updated representation of one's extemal environment. The unfamiliruity can also lie in the combination of the present perceptual environment (or intemal environment) and present motivation; for instance, if one has never been *in this area* and become *tired* before or that an *enemy* has never tumed up *here* before. Finally the 'novel aspect' of the situation can be related to recurring changes, for which one may also to some extent be prepared; only one does not know exactly when the change will occur or exactly what form it will take. (For instance, a particular

---

[40] In this chapter I am concerned with a *systems exp/anation* of *immediate planning,* in chapter 11.5 with a *systems explanation* of *anticipatory planning,* and in chapters 10 and 11 I deal with an *evo/utionary explanation* of both immediate and anticipatory planning.

Nettle-hedge withers.) Or the situation may concern more extreme, unusual and unexpected happenings. (For instance, a Thom-treethat has fallen down.)

3. Immediate Planning Motivation: The General Organization ofan 1-Creature's Planning ofAction

So, these are the kinds of situations in which an 1-Creature plans. Why? What makes the individual do so? Fora given individual anda given point in time one can ask: Why does it *plan* right now - and plan to salve this particular problem - and does not *do something else,* think of something else, engage in imagination or engage in a current problem by trying things out? This is not a trivial question in a system that is capable of several kinds of activities, external and intemal, and where certain activities occur at the expense of certain others (partly due to the system's limited resources). In soine way, the planning of action must be given a place amongst the system's other (behavioural and intemal) activities. Toere must be some kind of overall regulation of an individual's planning activity, for 'turning on' and 'switching off' planning.[41]

The basic solution to this design problem - the problem of immediate planning motivation - turns out to be simple in the case of the 1-Creatures. Fundamentally, remember, an 1-Creature is *a perceptual-behavioural system* with motivational mechanisms that ensure that it engages in its various domains of interest in a sensible way according to the principle of one-interest-at-a-time (see p.28). The primary motivations, hunger, fear, cold and tiredness, guarantee that an 1-Creature will devote its time in an adequate way to the corresponding problems and interests. And when it has no serious problem, it will be motivated either for exploration and play or for withdrawal and resting. For its primary interests an 1-Creature is equipped with a substantial amount of routines and programs that it acts upon in habitual situations. In situations that it interprets as novel, it may use its capacity for behavioural trial-and-error.[42] (See figure 20.)

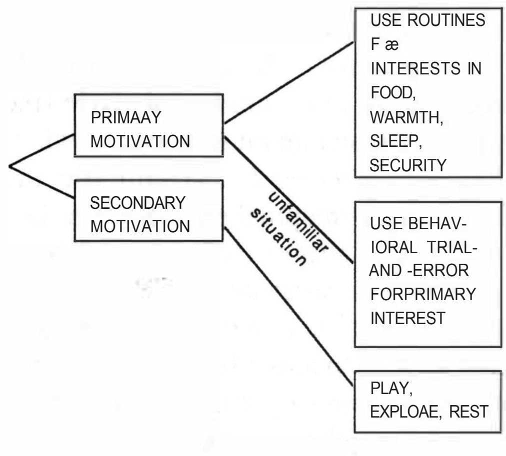


Figure 20. An 1-Creature's engagements and basic motivational organization as a perceptual-behavioural system

41 We are not interested in a system that can and does do *nothing but* plan (its actions) - however good. Rather, we want a system capable of planning and of using this capability in a sensible way together with its *other* behavioural and cognitive capabilities. We want *a planning* system, nota *planning-system.*

42 And this capability is used also in exploration and play.

Now, a simple method for 'switching on' planning in an 1-Creature is to have planning- that is, *internalgenerating-and-testing* of actions - turn on, when the individual

(1) has *aprimary motivation* and corresponding problem *and*

(2) as a pure perceiver-and-behaver would engage in *behavioural trial-and-error* to solve this problem.

This means that there is really very little we have to change concerning the system's motivational organization. We only replace *external* trial-and-error in unfamiliar situations where the individual has a serious problem with *internal* trial-and-error. (We do not, however, replace the behavioural trial-and-error involved in *exploration andplay.)* (See figure 21.)

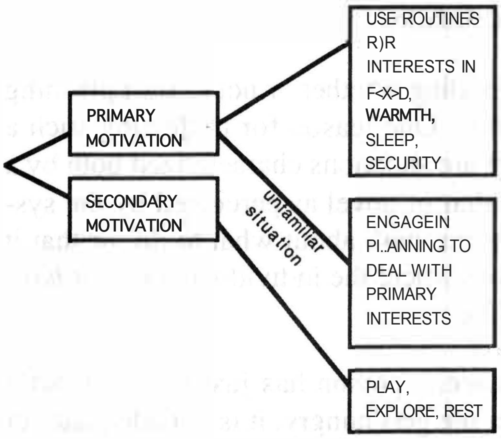


Figure 21. An 1-Creature's engagements and basic motivational organization as a perceptual-behavioural and *planning* system

The next problem is to decide when to interrupt or quit the ongoing planning of action. A first criterion for this is the following: 'when a plan, specified to the level of preformed programs or routines, has been produced and accepted'. When this happens the system gets out of the imagination mode and into the perceiving-and-behaving mode. Second: 'when the situation - i.e., the individual's motivational state and/or perceptual environment - changes in a decisive manner'. (For instance, an individual is tired and is planning where to go to get some sleep, when freezing takes over the tiredness. Or a hungry individual, enclosed and without food, plans how to get out, when suddenly a door is opened. Or, an individual is hungry and is thinking about how to getto a Berry-tree, when suddenly an enemy turns up.)Third: 'when the motivation gets very *strong* and the system is not succeessful in finding a plan for action'. (For instance, an extremely hungry, tired or freezing individual stops to think and reason about an unfamiliar and difficult situation and starts to do something.)[43]

43what we do not find in an I-Creature is engaging in planning when it has only a secondary motivation. There is no planning involved in play and exploration. Here it does not think through consequences and relate actions to one another before acting, but it just acts. Also, an I-Creature will not engage in planning just because it is bored, nor because it is anxious and seeks tranquility.

In a system with only these coarse methods for planning regulation the following obtains. *Whenever* the system has a serious problem and interprets the situation as novel, it will necessarily start planning. To try to produce a plan for action always has priority over starting to do something. Also the criteria for interrupting planning are coarse and give *priority* to the continuation of planning. Put crudely, planning is to be interrupted or stopped only in extreme cases.

Now, these ways for regulating planning work well enough for an I-Creature. But we can imagine a more sophisticated regulation of immediate planning, due to some added capacities for *evaluating*

• whether this is a situation in which it is adequate to sit down and think, and try to make a plan for action, and

• whether it is adequate to *continue* planning in this situation.

This would be a planner with more sensitivity in deciding whether or not to start planning and whether to continue or quit planning for action. One reason for preferring such a more sophisticated planning regulation is that there are situations characterized both by a serious motivation for acting and by being unfamiliar or novel as perceived by the system, where it is yet preferable that it does not start to think about what to do, or that it does not continue to do so. This is true for situations where the individual does *not have enough information* for making an adequate plan, like:

• If the information it has is outdated. (For instance, a person has just returned, after many years, to a town where she once lived. When she gets hungry, it is not adequate for her to spend too much time on planning where to go and on how to go to a particular place she has chosen until she somehow finds out which restaurants are still in business, what public transportation is available, etc. Or take the hungry Berry-Creature that is in an area where it has not been for a long time. Here it will not be adequate to spend lots of time on thinking about where to go before knowing about what paths and obstacles there are around.)

• If relevant changes are quick (and unforeseeable). (For instance, if you are sailing and want to sail a long distance from one place to another, it is of no use to make a whole plan for this (and particularly not to plan in detail) if the weather circumstances, the wind conditions in particular, are quite unstable. Or similarly, if a Berry-Creature is on its way to some eating place and the wind is shifting, there is no reason for it to try to plan (in advance) how to move to avoid walking in headwind.)

• If relevant aspects are difficult to survey. (This may be a combination of the two kinds of situations described above.)

• If the situation is *too* different or divergent from what one is used to and therefore difficult to make sense of. (For instance, if an individual is placed in a spaceship (alone) and gets hungry. Or if a hungry Berry-Creature enters some human environment. These are situations where it is hardly useful to sit down and reason, at least not long, but rather try to *find out* - by acting and testing - what can be done, what is edible, how one can get hold of it, etc.)[44]

It is preferable with a planner that is able to recognize such things, and to accordingly refrain from or to stop planning under these circumstances. In particular, some flexibility in shifting between and coordinating *internat* and *external* search (planning and behavioural trial-and-error) is desirable, as well as a capacity to *plan external search* for information to use *infurther* planning.

Of course, the 'stopping criterion' above of 'having produced an entire plan, specified to the level of routines or pre-formed programs' is coarse. At times it is preferable just to produce a plan where at least parts of it are more *sketchy* - like if it is the case that some information-search shall be involved before continuing to plan. At other times, on the contrary, the production of *one entire* plan should not be sufficient to start action; as in example (1.2) where the planner first makes a plan for how to get to a particular site but then considers another way to get there. This is an instance of planning that is not so much a search for a single solution but rather a generating of different solutions and then choosing between them in order to obtain a hetter, or the best, solution. Here the criterion for terminating planning must be something else, something like a 'level of aspiration' (see Simon, 1969). It can indeed be important that some such criterion is used because a problem with the kind of planning just mentioned can be that the planner *dwells too long* on the generating and search for a hetter or the best solution. This problem will not become critical, though, as long as the problem space is not too complex and the different possible alternatives are not too many.

The point is that we want a planner that uses its capacity for planning primarily in such situations and in such ways that it really *helps,* by reducing <langer, increasing speed, reducing effort spent, etc., and we want it to work in a sensible way *together* with routines, preformed programs and behavioural trial-and-error. Indeed, it is in a planner with a more sophisticated planning regulation that we can first speak of a system that is capable of *controlling its impulses to act,* or to *decide* to plan *instead* of acting or vice versa. It is also first here that we can find a system with a potential for improving its planning; *learning* from planning and action experiences. It can learn about in which kinds of situations it is of value to plan or not to plan; learn hetter when and how to plan its actions. (This is really the fourth step in Dennett's generating-and-testing ladder. See p.108.) But note also that such a sophisticated regulation requires a system with some *knowledge of its own knowledge* and its own cognitive and behavioural capacities and of different kinds of problem situations.

In conclusion, even though there are features of a more sophisticated capacity for immediate planning that may not be so easy to implement, it is not too difficult to fit basic immediate planning activity in a perceptual-behavioural system. And *in particular* nothing- and this is worth emphasizing - has to change concerning the motivation of the general

---

[44] Humans may of course often use some form of communication for obtaining information, for instance, by looking at a map, asking someone, asking at the tourist office (like in the case where the person returned to an old home town) listening to the weather forecast (if one plans to go sailing), etc. But during other circumstances, and always in non-social creatures, the only way to get information is to act, move, investigate, explore and try out various alternatives for oneself.

content and direction of the planning activity at a particular point in time. The general content, the goal and domain of the plan, is directly given by a *current* interest and motivation for acting. The same hunger that motivates behaviour to deal with the food interest can also make sure that the individual deals with this interest in imagination and thought. The basic motivation-system functions as a *general* filter or focus for distributing resources - *external as well as internal* - between interests (see figure 22).

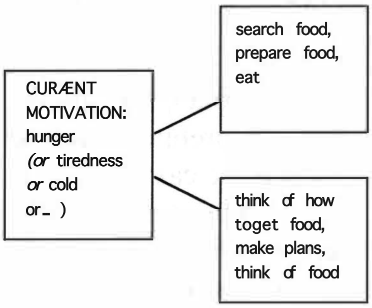


Figure 22. 'What to do now?' in an 1-Creature

4. Immediate Planning Knowledge

The question to be addressed in this section is the following: Given that a system in some situations is motivated to plan, how does it go about doing so? I have proposed a general description of the planning of action as internalgenerating-and-testing of plan elements and suggested some ways in which to limit and direct the generation of those plan elements. But in order to plan adequately, it is not sufficient to be motivated for planning at the right moments and to possess some mechanisms for the generating-and-testing of plans. It is also essential that the planner *has adequate knowledge,* organized in a sensible way. And so even though it is not possible to go into details, I want to discuss some of the more important *kinds of knowledge* that an 1-Creature has; under the headings of categorization knowlede, spatial knowledge, knowledge of change and self-knowledge.

4.1. Categorization knowledge

I have repeatedly retumed to the idea that perception - and memory - is hierarchically organized. The perceptual apparatus *constructs perceptions* that are composites of and invariants over *lower-order* perceptions. Situations, things, events and actions, are constantly perceived at various levels of abstraction. Meaningful chunks - configurations, objects, motion, events, etc. - are abstracted and represented at various levels.

Toere are two important observations to make in this context: (1) The synthesizing of sets of lower-level perceptions into more complex higher-order perceptions makes it possible to classify or *categorize* perceptual sets; for instance, a cluster of smell, form, texture, etc., as 'berries' or certain finger movements as 'grasping' (and do so even when some aspects from the cluster are missing). Different clusters or sets of lower-order perceptions *instantiate* a higher-order category. (2) Conversely one can go *from* the higher-order complex or invariant and find different instantiations of it, that ts, spell out details on the lower levels.

Some categorization knowledge and, most importantly, the inclination to *construct* such knowledge is built into the perceptual apparatus. It is categorization knowledge that makes it possible to classify groups of configurations, movements, operations, etc., and determine that they instantiate a more complex entity; for instance that jump over, walk round, get under, all belong to the category 'getting past an obstacle' or that several different kinds of berries instantiate the category berry or food. Conversely, categorization knowledge makes it possible to find instantiations of a category; that 'to get something down' can be 'throw something at it', 'push it over', 'drag it down' or 'lift it down', and so on. Categorization knowledge is of course involved in *generalisation* and *specialisation* of potential plan-elements, like programs and bits of programs. These are two processes that, according to the discussion on page 104, are crucial for planning.

4.2. Spatial knowledge

The 1-Creature's plans of action all relate to locomation - more or less centrally. They plan to do things at certain sites and to go to certain places. A spatial representation of the external environment - that is, a representation of where certain things are (and where certain things can be done) and of where such sites of interest are located in relation to each other-is a most fundamental aspect of an 1-Creature's internal environment.

A notion that is frequently used when discussing spatial knowledge in both ethology and in psychology - human and animal - is that of a *cognitive map.* There is not much agreement about the notion. Yet there seems to exist some points of relative consensus that I will tty to relate.45 By saying that a system has - and makesuse of - a cognitive map, one means the following:

• The system has an ability to *transform* information about a region that has been *sequentially* acquired in a series of successive experiences into a *simultaneous* cognitive structure in which the distance and direction between various successively experienced objects are indicated. In other words, it can read sequentially acquired information into a format where all path segments are equally available, whether they have been previously taken or not.46

• The system can *update* spatial knowledge in an *inferential* way, independently of present-time perception.(Thus it can handle, for instance, situations where all the spatial relations cannot be seen, because the space is too big or too complex.)

• The system *operates* on spatial knowledge in a way that is analogous to viewing the space from different view points.

For example, an 1-Creature may for instance walk around at a site A, notice that a Nettlehedge has disappeared and then go to site B and notice a missing Nettle-hedge there as well, and then later imagine both these sites simultaneously. It may upda e the path between the sites (as in example I.2) even though it can not see it. It operates on its knowledge *as if* it saw space from different stationary points.

45cf. Thinus-Blanc, 1987; Vauclair, 1987.

46This is in contrast to a structure of list- or register-format, where successive experiences are only sequentially stored.

In discussions of cognitive maps, one often emphasizes their 'as-if'-quality. They are *not maps* but *'like maps'.* My view is that the most misleading aspect of the analogy is that a prototypical map is static. It is constructed- and then it is there. All the dynamics that are possible are in our scanning it, exploring it, etc. But *cognitive maps* are often *continually up-dated.* (Remember, they do not only represent relations between sites but also 'what is there' at the particular sites.) Furthermore, a cognitive map may interact with search- and exploration-processes in another way than a conventional map and supply some of the dynamics itself. This, I think, relates to a more general problem that we have when we think and talk of *mental representations.* We tend to think of them precisely as static so that all dynamics must come from *operating upon the representations.*[47]

4.3. Knowledge of change - of actions and events:

The I-Creature of course needs more than knowledge of the spatial layout of its environment. Its planning essentially involves use of knowledge of events and actions of different kinds. A planner must have some knowledge of *processes:* information about what can and does change in its environment and how. It needs to know about what types of events regularly follow certain other types of events and actions.

The ability to plan, it seems to me, must in a fundamental way be grounded in some realization that certain aspects of the environment can be *changed* or *moved around* in order to obtain certain goals. *Minima/ly-and* crucially- a planner must have some knowledge of processes that it can *bring about* or *prevent itself,* i.e., knowledge of the changes that involve its own actions: for instance, knowledge that if it hits some berrys, they fall down, that if it climbs into a tree, enemies cannot reach it, and so on. Of course such knowledge is incorporated in various kinds of memory perceptions - in particular in *program references.* A program for action contains, as we have seen, representations of condition, precondition and outcome relations. Here are some examples: 'If you have a stick (precondition) you can use this to knock down berries (outcome)'. 'When there are some berries that grow high up (condition) you can use a *long* stick in order to knock down some berries (outcome)'. 'You can wrench a stick until it becomes loose (outcome)'. 'If it is windy and cold and you go down into a cave (condition), you do not suffer from the wind and cold anymore (outcome)'. Furthermore, a planner needs knowledge of certain changes and processes that do *not* involve and *cannot* be influenced by its actions - changes that it cannot do anything about but that are yet relevant for its actions and planning. An 1-Creature, for instance, uses knowledge of obstacles that have tumed up and disappeared, and of wind conditions. Such events may be conditions or preconditions for actions, as well as disturbances for certain actions.

Processes or changes can take more or less time. Within seconds and minutes the wind changes, it starts to rain and an individual seeks shelter, an enemy turns up, an individual picks some berries, and so on. An 1-Creature perceives these processes (events and ac-

47Several researchers, in particular some investigating imagination, try to change this way of thinking in that they suggest that expressions like 'to rotate an image', 'to transform an image', 'to scan, inspect or examine an image' should be replaced with 'to imagine a rotation', 'to imagine a transformation' and 'to imagine scanning, inspecting or examining'. (Compare the representation of *walking* by taking a paper doll and step by step make it walk or by pushing a walking button that turns on a film with a walking doll.) (See, for instance, Shephard, 1984; Freyd, 1987.)

tions) and can coITespondingly store and retrieve these perceptions as memory perceptions. And so, for instance, an I-Creature may perceive and imagine the wind blowing and associate this to trees and trees blowing over, if it has perceived this regularly (or if this has occuITed in a rare but important experience). But there are also *long term* changes; the wind chartging over a day, light getting dark, fluctuations over periods of time, and so on. Of such changes, however, an 1-Creature has no conceptions or representations. All its activities at a given point in time, remember, are entirely regulated by a present interest anda motivation for acting- and only knowledge of those changes that fall within the reach of one interest and need can be *used* by an 1-Creature. It is of *no interest* to an I-Creature to imagine, predict and represent processes that cannot have any influence on their behaviour or on their planning. Therefore it is only those reference perceptions that occur within the temporal frames of one interest that are ever *re-presented* by 1-Creatures.

4.4. Self-knowledge?

Now we come to a delicate issue. What knowledge does an 1-Creature have of *itself?* Does it have a *representation* of itself and, if so, of what kind? The reason that I bring up this question is that it is often more or less explicitly assumed that self-knowledge is a necessary requirement for a system capable of planning its own actions. I want to suggest that this may not be the case. I am aware, though, of the enormous complexity of this subject, and the following remarks are not more than tentative proposals.

In discussions of a *basic* self-representation or of the *origins* of self-representation, *the ability to recognize oneself in a mirror* is sometimes suggested as a sign for such basic self-representation. (See, for instance, Gardner, 1982, 1983; Gallup, 1977.) Let us take this as a hallmark of self-representation. What must occur when an individual recognizes itself in a mirror? Something like the following, I suggest: the individual somehow *identi.fies* what it perceives in the mirror with its own body as perceived directly. It does this although these two perceptions relate to entities perceived as (and represented as) *spatia/ly* clearly *separated.* The miITor image is there, I am here. Furthermore - and this seems crucial to me - there is not only the understanding of *onese/f as an object that can be perceived,* but also of *onese/f as perceiving. I* look upon *myse/f in* the mirror. I am an object - amongst other objects - but *a perceiving object.*

Now, is this ability something that is required for an 1-Creature' s capacity of planning its actions? No, in principle the I-Creatures can do without this. The concept of oneself as an object that can be perceived and as a perceiver, I suggest, is first necessary in order to plan actions where one takes other agents and perceivers into account as agents and perceivers; that is, in order to plan certain *social* actions and for social behaviour. In these situations also the third element of a full self-representation is involved, namely, that 'I ama perceiver amongst *other perceivers'.*

On the other hand, some of the competences that an 1-Creature has can be of use in the development of a self-representation:

(1) There seems to be a common core in the ability to identify, say, the twig that I see before me with the twig, broken off from the tree and with its leaves peeled-off, that I imagine and plan to use as a tool; thusly relating a *present perception* and an *imagined* reference perception (a capability that the 1-Creatures have), and the ability to identify what is

in the mirror with me standing here; thusly relating *two present perceptions* to each other. *Both* abilities involve the identification of objects simultaneously present but represented. as spatially sepaiated.

(2) The 1-Creatures' ability and inclination for *exploration* certainly can be useful for the development of a self-representation, if they use it for approaching their own bodies as objects to explore. (See Lorenz, 1978, p.210.)

An 1-Creature may thus have abilities that are *precursors* for developing *some self-representation.* It might represent itself and others (1) as objects to be percevied and (2) as perceiving objects or perceivers. But what would *necessarily* be lacking in an 1-Creature's self-representation and representation of other 'selves' is an ascription of needs, desires and motives, etc., to these selves. The ability to represent non-present, non-perceived needs, desires and motives, recall, is one crucial difference between the 1-Creatures and the A-Creatures.[48]

In summary, a self-representation is *possible* in an I-Creature or in an immediate planner, but it is neither *required* nor would it be of any *use* for an 1-Creature as pianoer or as a perceiver-and-behaver. One reason why a self-representation is often listed as a necessary ingredient for a planning ability can be that one has social planning creatures in mind. (And it is undoubtedly the case that more sophisticated planners in the biological realm - those first thought of as planning creatures - are also social creatures.)

5. Conclusion

In this chapter I have given a characterization of an immediate pianoer. As I have already pointed out, this is not to be taken as an abstract description of *any possible* immediate pianoer. Instead, it is a description that includes several assumptions about what seems to be true of what I take to be immediate planners in nature. For instance, other kinds of interest than those that the 1-Creatures have are imaginable, planners without a capacity for play are possible, other kinds of motivation-systems than the one I describe are possible, and so on.

My prime aim remember is to gain a hetter understanding of natura! biological planners. This remains true in the next chapter where I will characterize an anticipatory planner and discuss the transition from immediate to anticipatory planning. That is, the suggestions that I make are, in the first place, hypotheses about a factual development. Secondarily, however, such proposals might of course also be helpful if one wants to *construct* a planner, in particular, an anticipatory planner.

48Social interactions in creatures with such limited self-representations must also be limited. They may handle 'overt signs' of motives in another agent where an enemy with open eyes implies a threat but not one with closed eyes, for instance. But, they can have no conception of ***deception*** where the enemy is closing its eyes to fool you. (Luckily the 1-Creatures enemies are not quite so elever.)

9. AN ANTICIPATORY PLANNER

1. Introduction: The A-Creatures; their Planning and a Characterization of them as Anticipatory Planners

A long time passes by. We forget about the 1-Creatures, but then one day we take a look into their world. We see some creatures with the same outer appearances. After some time, though, we realize that they differ from the 1-Creatures. We notice some differences in their behaviour:

• A-Creatures sleep in Umbrella-trees as did 1-Creatures but they manipulate these trees in a certain way. They take very thin pliable and long stringlike twigs from Bast-bushes and use them to fasten down the branches of an Umbrella-tree, in a more vertical position. You could say that they are building a kind of night-shelter. We realize that this makes a difference during nights when it is very windy. It is not a very simple enterprise. Bast-bushes do not grow together but are scattered, and there are not many of these long twigs to pick from one bush. Furthermore, the fastening of the branches is difficult and takes time. And so, the construction of a night-shelter takes a few days. Under certain circumstances it can be done quicker: When it rains (which does not happen so often) these long twigs sprout, and for a short period there are thus enough twigs for building a night-shelter by using just one bush. After a night-shelter has been built in this way, the Berry-Creature uses that Umbrella-tree as its 'sleeping-tree'. After some time, however, the leaves of the tree start to wither, and they do not provide enough protection. Then it is time to build a new night-shelter.

• When an A-Creature goes to its Umbrella-tree, and often when it goes to a cave to get shelter from cold, it brings along a small amount of berries (which means that it does not have to leave its site to obtain food when it gets hungry).

• A-Creatures spend more time than the 1-Creatures in what seems as resting in various trees and in caves, and they spend a little less time in play and exploration.

• After having used two stones to crack a Shell-Berry, an 1-Creature (cf. p.65) will just leave the stones, and the risk is high that the stones will not be there the next time they are needed. But A-Creatures are sometimes observed digging a hole into which they put the stones and then cover it. The next time they come to the Shell-Berry-tree, they dig up the stones and use them again.

• We can observe some individuals doing certain things with Nettle-hedges, apart from cutting them down to pass them, as the 1-Creatures also do: They fasten a Nettle-twig onto another twig which functions as a handle, and then they bring this 'weapon' along when they pass 'a dangerous area'. If they encounter an enemy, they use this for hitting or throwing at the enemy.

In other aspects, the A-Creatures behave just as the 1-Creatures. They eat the same things, use the same tools and methods for getting food, go to caves when it is cold, flee from enemies, engage in exploration and play, etc. One can still subsume their behavioural engagements under the primary interests in dealing with food, sleep, shelter, threats (enemies mainly) and the secondary interests in exploration and play and in resting.

The differences between the A- and 1-Creatures, which are the most interesting in the present context and which are indeed a basis for the differences in behaviour, are *differences in their knowledge and mental activity.* Consider the following examples of planning in A-Creatures:

Example (A.l): 'Planfor sleep andfood'

M(- ; Boredom) PE(is at a Shell-Berry-tree, has been eating Shell-Berries and is now satiated).

Thinks about what to do now: Go and explore the area on the other side of the hill? Yes. But it will soon get dark and I will get tired. Therefore I will not go too far. Go to the left or the right of the ditch? To the right is better, then I will be near my Umbrella-tree when I get tired. I can also take some berries along from the Good-Berry-tree further down when passing.

Analysis: This individual is considering*future* interests of food and sleep as well as a current interest of exploration. (An 1-Creature in a similar situation would just go exploring and playing.)

Example (A.2): 'Planfor de/ense'

M(- ; Boredom, curiosity) PE(is in at a site where there are several Shell-Berry-trees, resting in a Shell-Berry-tree).

Imagines the following situation: I am in the Good-Berry-area, and have found some very big Good-Berries, but then an enemy turns up. So I will take a Nettle-weapon that I have brought along, hit the enemy and then I flee"but"no, I can do this instead: before I start eating the berries, I will take branches from the Nettle-hedge and *cover* the path with them. And then, when the enemy comes, it cannot pass and willjust turn away.

Illustration 8. Dealing with a potential threat

Analysis: Here an imaginative individual invents a method and constructs a plan for dealing with *apotential threat.* (An 1-Creature in this situation would engage in some playing.)

Example (A.3): 'Planfor she/ter andfood'

M(Freezing; Anxiety) PE(is in a High-tree, resting, and is now getting cold)

Thinks of what cave to go to and is at first uncertain: I don't know these surroundings well, but wait, there is a cave nearby further down this path. Ok, but as there is no food near that cave, I will then first go to the Red-Berry-tree on the other side of this ditch to get some food, because if it stays windy, I want to stay for some time in the cave and not have to get out.

Analysis: The individual considers a potential food interest as well as the current interest in finding a shelter when it plans what to do. In relation to the latter, the path it plans to take means that it will make *a detour.* (An 1-Creature would in this situation concentrate on getting to a shelter as soon as possible and not take any food along.)

Example (A.4): 'Plan what to do, andplanforfood and de/ense'

M(-; Boredom, restlessness) PE(is in a cave,where it went earlier because it was frozen).

Thinks about *what to do next:* Maybe I shall go and get some food. It has been a while since I last ate, and only some Red-Berries, so I will probably get hungry soon. Where shall I go then? There is a Red-Berry-tree very near. But I would rather have some Shell-Berries. A Shell-Berry-tree... there is none around, well there is the one, but I would have to walk around the big ditch. Or... maybe there is no water in the ditch now. I could check that first. But wait - I am really not that far from those *Good-Berry-trees.* No, but again, there are ditches, there is no straight way from here, that is, if the ditch along the Kay-bushes is not dry... - I shall go there and check it. Then, if I can pass there, I will first make a Nettle-weapon and then go into the area and have some Good-Berries. If I cannot, then I will go back and have some Red-Berries here, or rather from the Red-Berry-tree that I will pass on the way.

Analysis: Note this individual thinks about food before it is hungry and also of a potential threat by an enemy if it decides to enter the Good-Berry-area. (An 1-Creature in a similar situation would go exploring the environment.)

Continuation ofexample (A.4):

The individual acts upon the plan and goes to eat Good-Berries.

M(Hunger; Anxiety) PE(is just about to start to eat Good-Berries)

Before it starts to eat, it thinks the following: If an enemy turns up now, what shall I do then? Ok, I have this weapon, so I could hit it and stop it for some time. And there may be time toget out from this area and to escape up in a tree... But the Good-Berries will be lost...

Analysis: Assume this is the same individual as in example (A.2) above. It may then *re-trieve* the plan about blocking the path to the tree with nettles and try to execute it. That is, it first makesa plan fora potential threat, a plan 'in case' (to use a weapon). Then it retri-eves another plan, again for the potential threat, and sets off to act upon this plan (to build a trap). (An 1-Creature in the same situation would just eat and not consider potential threats or take precautions for one.)

Example (A.5): 'Plan what to do, plan to play'

M(Hunger; Curiosity) PE(is eating Shell-Berries)

The A-Creature is thinking of what to do now after it has eaten: Well, I will first put back the big stones into the hole and cover it. And then? What shall I do until it gets dark and it's time to go to sleep? I could go and climb in those trees over there. That seems difficult. If I could learn to climb those maybe another day I will then also be able to climb the Shell-Berry-tree further down the river that I failed to climb earlier. Or I can go and play with stones at the Plumb-Berry-site and take the opportunity to check the Bast-bush on the way to see if there is any material for making a new night-shelter another day, but no - that can wait really.

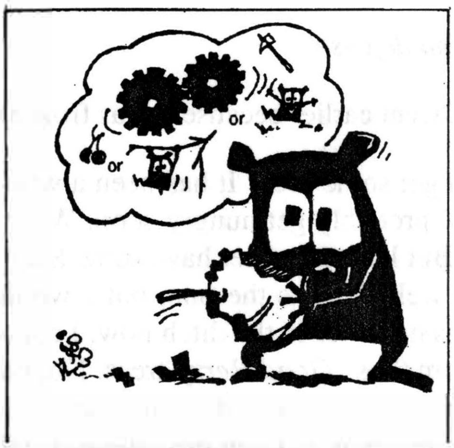


Illustration 9. What to do next

Analysis: This is a situation where an 1-Creature would not engage in planning or imagination at all but just eat its Shell-Berries. Furthermore, an 1-Creature would never *plan to play* in the way illustrated in this example.

Example (A.6): 'Plan when to fix a sleeping place'

M(-, Anxiety) PE(has woken up in the Umbrella-tree, resting, sees that the leaves of the tree are beginning to wither)

The A-Creature thinks of the following: *When* shall I fix a new night-shelter? It was difficult last time and took a long time, and then just a few days later there was a heavy rain, and lots of sprouts appeared on the Bast-bushes. Now this time, then, I should maybe wait until the next time it rains? I can wait a couple of days at least. Although if it gets really windy before that, I had hetter start fixing something sooner of course.

Analysis: This individual thinks about fixing a night-shelter and about *when* to do this, without being either cold or tired. (An 1-Creature could not be in this kind of situation, as 1-Creatures do not build night-shelters in this way.)

A-Creatures are obviously *anticipatory planners*. They plan actions that do not only relate to a current interest in getting food, getting shelter or escaping from a threat, etc. - but also to potential, anticipated needs and problems that they think of. An A-Creature, as we have seen, can for instance plan actions like (1) taking a 'weapon' along when passing a certain area in order to be able to protect itself from a threatening enemy *if* one happens to turn up, (2) taking a certain path in order to be near an Umbrella-tree *when* it gets tired, or (3) fixing a night-shelter now, or today, sothat it will later have protection against the wind *when* it gets windy and the creature will get cold, and so on.

An *1-Creature* approaches its environment in the following way: How can this environment be transformed into one that solves this *current problem* and satisfies this *current need?* But an *A-Creature* also 'asks': What will happen in this environment, and what needs and problems can then be foreseen? And what can be done about these problems- now or later? How can the environment be transformed in relation to those potential needs and problems?

An A-Creature may *considerfuture goals andfuture actions* that it is *not* at all *interested* in obtaining or performing *now,* like in example (A.6) where the individual is not interested in sleeping now but plans some arrangements in order to be able to sleep well later on. And it can think of and plan an action that is *not possible* to perform *now,* as the relevant problem situation does not exist- and may even never come to exist-like in example (A.2) where the A-Creature planshow to deal with a potential enemy that could turn up at a certain site.

If an 1-Creature, as an immediate planner, is a problem *solver,* an A-Creature, as an anticipatory planner, is moreover a problem *searcher* and even problem *constructer.* The questions I now pose are the following: How can an 1-Creature develop into an A-Creature? How can we change a system that is merely competent of and motivated for· immediate planning into a system that is also capable of and motivated for anticipatory planning? And what difficulties will we encounter in achieving this? The next sections are an attempt to answer these questions. Toere I try to give a systems explanation and understanding of an A-Creature.

2. When Does Anticipatory Planning Occur?

The first question to be answered about the A-Creatures is under what kind of circumstances A-Creatures engage in constructing plans for actions. Some of the situations where an A-Creature engages in planning are situations where an I-Creature would do this as well. Example (A.3), for instance, illustrates this. The individual has a primary motivation (being cold) and perceives some unfamiliarity in the situation. However, the difference in this example lies in the fact that the A-Creature does not *only* take this *current* primary interest into consideration. (And in this example, an I-Creature and an A-Creature would come up with different plans.) And of course there may also be situations where an A-Creature will engage in 'pure' *immediate planning* just as the I-Creature, like in very pressing situations with strong hunger or fear or tiredness (even though A-Creatures may, due to their anticipatory planning, eliminate some of the problems that I-Creatures solve by immediate planning).

But, there are also situations where only an A-Creature, but not an I-Creature, engages in planning, like in most of the examples, pp.122-124. These are situations, namely, where the planning is not directed by a current serious need and corresponding motivation for acting; situations where an I-Creature would either engage in behavioural problem solving or in exploring and playing around. It is not as easy to give a general characterization of 'an anticipatory planning situation' as it is for 'an immediate planning situation'.

One thing that can be noted is that, *on the whole,* A-Creatures *plan more* than 1-Creatures: They spend more time in an imagination mode thinking about how and when to obtain various goals and do certain things and on what things to do and when.

3. Anticipatory Planning Knowledge[49]

Are there any *new kinds of knowledge* that an A-Creature has and must have, and that an 1-Creature is lacking? Consider the following sample of ideas or thoughts that an A-Creature may have in contrast to an 1-Creature.

• The idea of probably getting tired and need to sleep before getting back from a certain site (cf. example (A.1))

• The thought that if it gets very windy today and it does not fix its Umbrella-tree, it may freeze when going to sleep (cf. example (A.6))

• Thoughts about *what interest* to engage in next (what to do next) or what to do *after* it has dealt with a *current interest* (as in example (A.4) and (A.5))

• The idea that this could be of use later when or if this need turns up, (as in example (A.5) where it saves the stone-tools)

• Ideas such as 'this is an opportunity', 'to miss a chance', 'not to waste this chance' - where these ideas relate to non-immediate needs. (This, for instance, can happen in relation to example (A.6).) Here the individual plans to take the opportunity next time it rains to build a new night-shelter. And so, if it rains, but the individual for some reason does not get the material for the night-shelter, it might conceive of this as 'missing a chance'.)

Basing upon these examples, you may realize that a crucial difference between A-Creatures and I-Creatures is the following. *Both* I-Creatures and A-Creatures - as immediate and anticipatory planners respectively- are capable of considering not only their *present* perceptual environment but also other *possible* perceptual *environments.* The 1-Creature, however, only considers such *possible environments* in relation to an *unchanged motivational state,* whereas the A-Creature, in contrast, ponders *possible environments* in relation to *current and possible motivational states* (see figure 23).

49Note that I discuss the issues of planning knowledge and planning motivation in the reverse order as compared to bow I ordered them when I discussed immediate planning. The reason for this is that it is useful to refer to anticipatory planning knowledge when discussing anticipatory planning motivation.

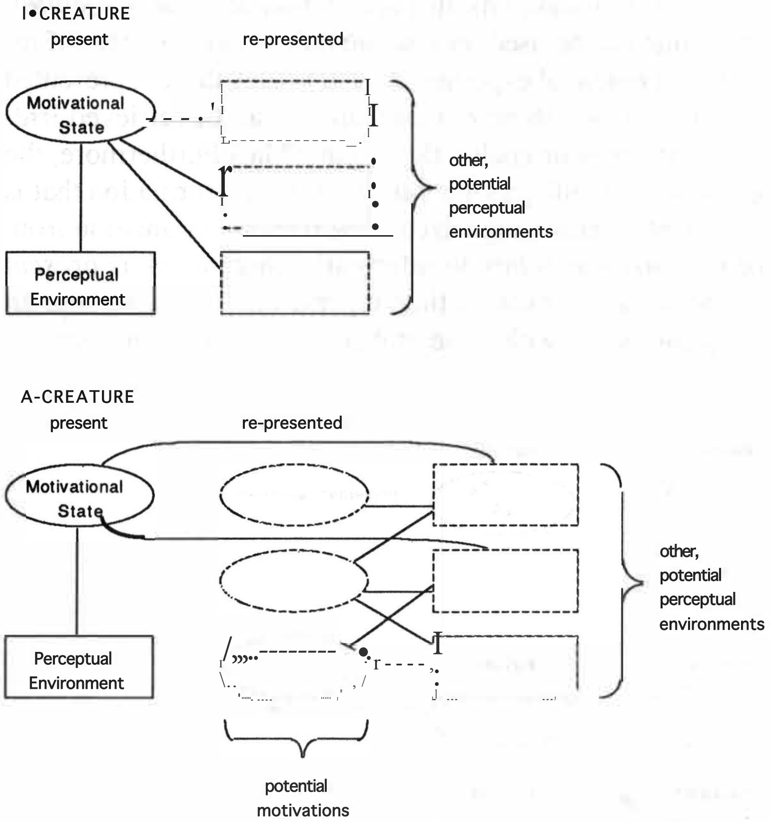


Figure 23. What 1-Creatures and A-Creatures, respectively, can think about

In the *1-Creature,* as we have seen (figure 22, p.116), its *current motivational state* is the general filter for directing attention and other resources, first, in that it directs *perception and behaviour,* but also in that it directs or channels the creature's *imaginative activities,* in particular its making plans.[50] But motivational states are not elements that figure in the *internat environment* of an 1-Creature. *Environmental* states can be imaginatively reproduced or *re-presented* by 1-Creatures, but it is only in *A-Creatures* that *motivational* states - needs, feelings, moods, drives, etc. - are represented.

Thus, the A-Creatures indeed have new kinds of knowledge. They, but not the 1-Creatures, have *knowledge of motivational states;* of bow they develop and what environmental states and changes they are influenced by. In contrast to an 1-Creature, an A-Creature must have knowledge of states and changes that go beyond the duration of one motivational state and be able to *conceive of* them independently of its present motivational state and interest.

These new kinds of knowledge, I propose, are organized around a *time conception.* The A-Creatures, but not the 1-Creatures, have what I call *an explicit time representation.* This *is a new cognitive structure* used for the organization of an anticipatory planner's knowledge and behaviour.

---

[50] The current motivational state directs search for solutions to the individual's current problems in its *internal* as well as *external* environment.

By an explicit time conception or representation I mean a conception of some 'structure', with a *present*, a *past* and a *future*, that can be used for a *simultaneous organization* of re-presentations of present, previous and potential experiences; that relates these represented experiences to one another (and from which these representations can be retrieved irrespectively of their relation to present needs or goals). (See figure 24a.) Furthermore, the 'structure' is dynamic in a specific way: it enforces the transition from information that is at one time organized *as present* to information organized *as past* and the transition from information that is at one time organized *as future* to information organized *as present* (see figure 24b). An individual who has an explicit time representation *knows* that an event or action that it *predicts* or *plans* will - with some probability - become present and then past.

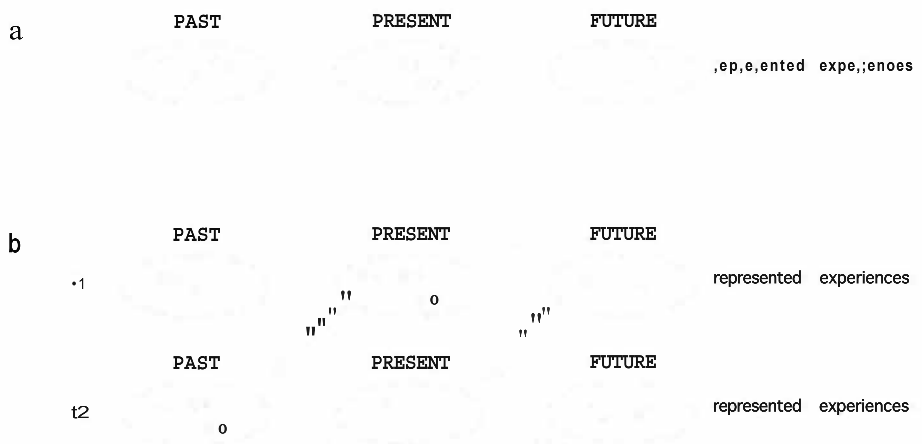


Figure 24. An explicit time representation

An explicit time representation can be further elaborated in many ways, with a more or less elaborate structure within the basic structure. Toere can be more or less sophisticated kinds of labeling or 'time-tagging', whereby past, present and potential events, actions and situations are related to each other. In other words, there can be a more or less sophisticated *episodic memory,* where episodic memory means 'time-tagged' knowledge of particular episodes and experiences. (A creature with an episodic memory does not just know things of the form 'if A happens, then B', 'when I do C, D happens' but also of the form 'at a particular time when I did C, D happened'. Knowledge of the first kind can be organized around or tied to knowledge of specific episodes.) (For the notion of 'episodic memory' see Tulving, 1972.)

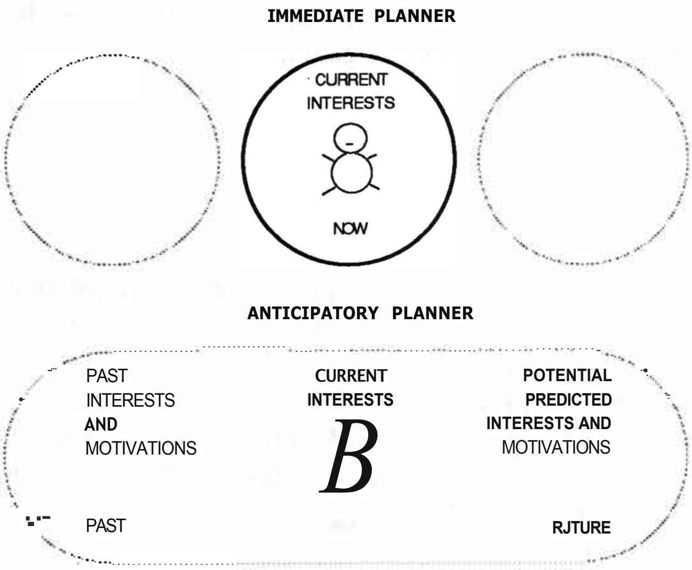
I will now discuss the issue of the explicit time representation in some more detail; first by relating it to spatial representation. It is essential for I-Creatures as well as for an A-Creatures to have *spatial knowledge.* It is important for them to know where things are and where things and places are in relation to each other. Therefore, it is important for them to be able to re-present *paths,* where a path basically consists of one thing or place A, another thing or place B, and a series of locations spatially connecting them. Path representations - in other words, (combinations of) memory perceptions of paths - are, of course, extensively used in the planning of *locomotion,* where the path representation is also endowed with directionality: one *moves from A,* the *start, to* B, the *goal.* But such a structure can also be used for representing actions where other aspects than locomotion are central. In these cases, let A and B stand for start and goal *states,* and let these be connected by a series of intermediate *actions and states* (see figure 25).

Figure 25. Path representation

Indeed an 1-Creature can in principle situate everything that at a particular moment of time is of interest for it somewhere *in space.* When it has a problem, the *goal state is not here* - but elsewhere. What it does is to seek to exchange its currently perceived space against another - to move somewhere, or to move things around, in order to change its perceptual environment. All that is of interest *is* or *takes place somewhere* in current space.[51] An A-Creature has to go beyond this and extend its organizational formats for its sphere of interests. In the following pages I will adress this issue.

For the 1-Creature, a present interest and what relates to this interest is all that is to be cared about. A current need is a *given* point of departure. It is self-evident and unequivocally *presented* to the creature. But in an anticipatory planner these seif-evident and indisputable 'horizons' around what is at a particular moment of time *relevant* or *important* are lost (see figure 26). The A-Creature makes plans that involve imagined motivational states, goals that are not to be reached now, problems that are only potential, etc. It consider interests, needs and motivations that are not *present* but must be *re-presented.* But then we must make the *representations* of these things part of the individual's intemal environment and connect them to what is already treated as worth caring about, of interest and real. This, I suggest, occurs not by placing these potential problems, events, states, motivational states etc. in space, but by using *a time representation.* Thus, the self-evident motivational point of departure and the center of importance for the *I-Creature* is in the A-Creature labelled - and represented - *as present,* and the ideas of *a future* and a *past* are introduced and tied to the present. In this way, an explicit time representation is an organizational format that makes possible a *simultaneous consideration and comparison* of presently experienced needs and future (and past) needs. Representations used for interpreting the present, and for remembering the past - thusly 'presenting' or 'making present' the past - are also used for predicting the future - thusly 'presenting' or 'making present' the future.

[51] It indeed seems that the method *closesest at hand* for representing what is not presently here but yet is believed in or aimed at and considered important probably is to situate it *somewhere else in space* - thusly connecting it to 'immediate reality', to here and now. There are same indications of this in utopian writings. It was not until late, namely, that the 'Utopias' were described as goals or ideals to be possibly obtained *later in time,* bot for quite some time they were concevied of as places *existing now* and to be *discovered.*



.Figure 26. Domain of interest and relevance for an immediate and an anticipatory planner respectively

It is by a *time representation* - not by space and motion representation - that the system relates present, remembered and predicted motivational states and that needs, feelings and other motivational states become elements of knowledge for the system, or, in other words, parts of the system' s internal environment. An *1-Creature* never thinks of a non-present motivational state. It has no stored perceptions of such states and therefore cannot retrieve and represent them. (The freezing I-Creature that has become warm again or the threatened I-Creature who has escaped its enemy leaves the problems of cold or threat behind in order to concentrate on the next motivation or interest domain.)

Again, for an I-Creature, a currentmotivational state (like present hunger) is the first and main *filter* for directing attention and resources in extemal as well as intemal activities - for directing what it does as well as what it thinks of. (Cf. p.116.) It cannot function in this way in an anticipatory planner, as it also engages in and devotes resources to potential interests and motivations. But we also cannot just take away this filter and let imagination run wild. We cannot just allow any possible interests to randomly enter the imaginational scene and consume resources. If this were to occur, we would not get planning but imagination anarchy. (Note that the seriousness of this is related to the number of motivations and interests a system has.) But the value of having an explicit time representation is that an individual can gain and organize some *knowledge* about interests and motivations and their unfolding in time and relate various relevant experiences to points in time, past and future. It can *remember* such states, situations and events and deal with them as real or serious even though they are non-present. With an explicit time representation there is a possibility for leaming and having knowledge of environmental changes that extend beyond one motivational state, and of bow various environmental changes can in:fluence needs and motivations. And so the A-Creature can, even if it is, for example, presently warm, in a useful way think of nightfall, of how it may get cold as it becomes windy, and remember bow, some days ago, it was very cold and of bow it solved this problem. Anticipatory planning, in contrast to immediate planning, requires cognitive structures of such a kind that they can be used for retrieving relevant memory perceptions at the right occasions, without hardwired, 'low-level' biological directions for doing so.

In sum, I suggest that, in the A-Creature and generally in an anticipatory planner, a time representation is introduced as a new principle for distributing an individual's resources and for organizing its behaviour and knowledge. The temporal organization of knowledge of motivational states and experiences tied to those states constitues a new imagination filter for the A-Creature. It is in this way that potential interests and problems are represented and can be dealt with in a meaningful way. (Cf. Bischof, 1985, pp.591-593.)

Once the step has been taken with a time structure for remembering and predicting experiences as belonging to a past and to a future reality respectively, and as something that can deserve attention, there is also a potential for an *extension* [52] of the time representation towards the past as well as the future. With such a time representation, furthermore, it is also likelythat an appreciation of some things in the past, and in particular in the future, as being *unknown* and *uncertain* will follow. Now such an appreciation in a creature with a *certain tolerancefor novelty and unfamiliarity* - indeed with a *need* for novel and unfamiliar information but *also* a *need* to perceive what is well-known, familiar and 'as usual' - may be difficult to handle for the system. (Cf. Bischof, 1985, pp.550-551.) (As speculation, this can be one source of a representation of time as 'object-like', as a path, a circle, a square, etc., to render the untangible and unknown time and future *some* tangibility.)

Toere is also a potential for extending a conception of time and rendering it more *abstract.* Even if the structure is, from the beginning, only filled with representations of*particular* events, processes, states or experiences, one may from this arrive at a conception of *the flow of time* that exists as such. Abstract time may be filled with *any* process. It is invariant over *particular* events, processes, etc., indeed to the extent that it is independent of any content at all (see figure 27).

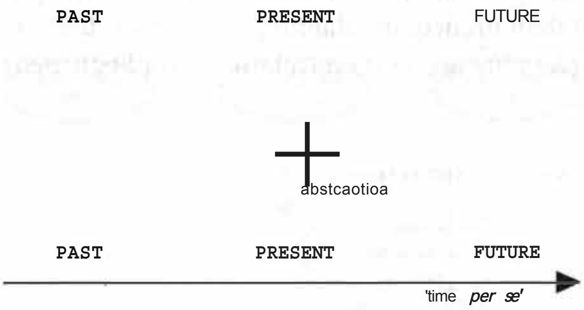


Figure 27. Abstract time or 'time per se'

So if we compare an anticipatory planner with an immediate planner, the anticipatory planner has some new *process knowledge* and a *concept of time.* On the other hand, it need not have any new *spatial knowledge.* What about *self-knowledge?* In the section on immediate planning, I suggested that a self-representation may not be required for imme-

[52] Possibly only as a 'by-product'.

diate planning. I will just make a few comments on this in relation to anticipatory planning and suggest that self-representation is also not in principle required for anticipatory planning. But I am just as tentative in this section as before.

Recall that by a self-representation I mean the following: first, the conception that *I am a perceiver* and can perceive objects (amongst other things); second, the conception of me as an object that can be perceived and *an object amongst other objects that can be perceived.* (These conceptions taken together imply that *I can perceive myself.)* The third aspect is the conception that *I* (and other objects) *can be perceived by others* - that I *am a perceiver amongst other perceivers,* other selves.

Does anticipatory planning necessarily involve a self-representation? What is new in the anticipatory planner is that it has an *explicit time representation,* by means of which it can compare and relate present and potential needs. This means that present and potential needs are treated as being *of one kind.* A present need or interest no longer plays the role as the self-evident reference point. It is not *alone* in its kind. Yet, the anticipatory planner as the immediate planner can well have *itself* as a self-evident reference point. It does not *have* to compare itself to other objects or other perceivers and thus conceive of itself as one amongst other objects or one amongst other perceivers (see figure 28).

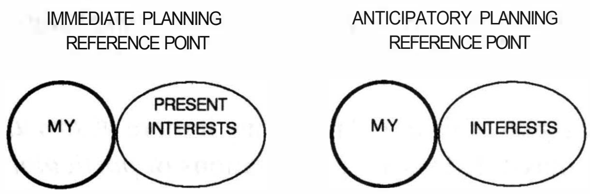


Figure 28: The seif as reference point in both immediate and anticipatory planning

A self-representation, I maintain, as I did in the section om immediate planning, is *required* first for certain social behaviour and for the planning of such behaviour. But anticipatory planning is no more intrinsically social than immediate planning. However, there is an interesting parallel between anticipatory planning and certain (relatively sophisticated) social behaviour (see figure 29).

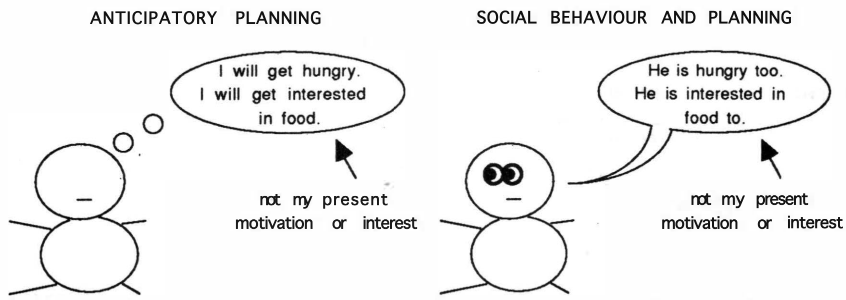


Figure 29. Relationships between an anticipatory planning ability and social competences

For various kinds of social behaviour the ability to consider and represent *motivations of other individuals* is required; for instance, for deception, for certain forms of cooperation, for acting out of compassion or in general for all instances of acting in order to meet someone's needs even though oneself does not have a similar need, etc. In the context of such behaviour, one must be able to infer and represent motivational states that are not directly *presented* like one's own current motivational state. But in *anticipatory planning* as well there (necessarily) are *representations* of motivational states involved - not of motivational states of others, but representations of one's own non-present, predicted ones.

Furthermore, a development of anticipatory planning in creatures might imply a pressure towards socialization. To wit, new activities, like shelter building, keeping some kind of food storage (which presuppose an ability for anticipatory planing), open up *new possibilities* for competition and cooperation and thus for the emergence or development of social structures for regulating or for dealing with this. But the influences can also go in the opposite direction, so that a development of social behaviour and corresponding conceptions could make possible or even imply pressure on development of anticipatory planning.

Whatever direction the influences take, there are indeed some social phenomena that are not possible until the (inter)acting individuals are endowed with an explicit time representation (phenomena like 'giving in order to possibly get something back', 'taking revenge, 'repaying', some forms of punishment, etc.). (Cf. Bischof-Kohler, 1985, pp.30-32.) But note that I am *not* saying that all social structuring requires or is tied to anticipatory planning. The phenomenon of social structures is definitely older than reflection and explicit thinking, and many forms of social regulation are not at all (or at most marginally) tied to the reflection or problem solving in individuals.

A final note: An immediate planner *can* (but does not have to) have a self-representation, and the same is true of an anticipatory planner. Yet an anticipatory planner *can* have a *richer* self-representation than an immediate planner can, in that it can represent itself and others not just as objects to perceive and as perceivers (or centers of perception), but also as centers of motivation and owners of motivational states. That is, the anticipatory planner may not only conceive of its own *perceiving* relative to someone else's perceiving, its own *physical appearance* relative to that of someone else and its own *abilities* relative to those of someone else, but also of its own *feelings, wants and needs,* etc. relative to those of someone else. In this way, the development of an anticipatory planning competence in an *immediate planner with a self-representation* could have effects on this self-representation in adding motivational features to it.

4. Anticipatory Ptanning Motivation

How is an A-Creature motivated to engage in anticipatory planning? Remember that we start out from an 1-Creature which has a well-functioning motivation-system that channels its resources in a strict and determined way according to the principle of one-primary-motivation-at-a-time (p.28). A current motivational state directs *all* the activities of an 1-Creature - behavioural as well as intemal (perceptual and imaginative) (cf. p.116). For instance, hunger motivates searching for food or eating, but if there is no ready suitable behavioural program to act upon, hunger instead motivates *thinking about* or planning how to go about finding food. It follows that *when* to plan and the *general content* of the planning activity is always given and unequivocal (see figure 30).

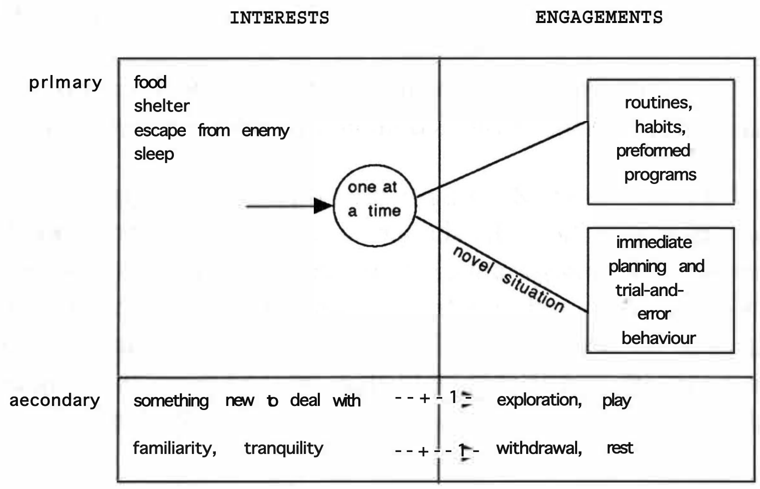


Figure 30. Motivational organization in an 1-Creature

How then can we fit anticipatory planning into the spectrum of the 1-Creatures activities? In other words, how can we make needs and problems that are *not* presently experienced but only imagined effective in directing imagination and thought - and behaviour? It is important to be clear about the problem. Departing from an 1-Creature's engagements or activities, we now have a new kind of activity to make room for. The A-Creature, at times, *spends resources* on thinking of potential interests and problems *at the expense of* devoting resources to other activities: to perception, behaviour, imagination and thinking that relates to current interests and to what takes place in its current environment. At a particular moment of time it engages in anticipatory planning at the expense of doing something else; like, for instance, searching for food, getting some sleep, paying attention to what is happening around, seeking a shelter, playing, exploring its environment, thinking about a current need, or just resting, etc.

It is, remember, the *motivation-system* that is the *evaluator* of what is important and worth caring about at a particular moment for the individual. The problem is: How can *potential* needs and interests and problems be *motivationally commensurab/e* with *present* ones? In dealing with this problem I will use the following point of departure: to try to make as much use as possible of the original motivation-system (that of an I-Creature), with its mechanisms for information processing and evaluation.

In the anticipatory planner, potential interests must be effective both in directing *imagination* and in directing certain *behaviour,* that is, in directing *internal* as well as *external* activities. In example (A.3), p.123, the planner *thinks both* of currently being cold and of potential hunger. Furthermore, bothofthese interests direct its *behaviour.* The planner 'takes a detour' toget some food and does not immediately act upon its being cold. In sum, *currently being cold* and *imagined, anticipated hunger simultaneously* influence and direct intemal and extemal activities. But the motivation-system in the 1-Creature, remember, works according to the principle of one-motivation-at-a-time. If it is the interest in food that is assigned priority by the motivation-system, this will *dominate.* The 1-Creature will in a determinate way engage in this interest, in acting, and, if required, in imagination and planning.

We realize that in the A-Creature we have to *give up this principle* as a basic principle of the motivation-system. Note thus that in order to change an I-Creature into an A-Creature it is *not sufficient* with a purely *cognitive* change (a change in knowledge and knowledge representation) while leaving the original basic *motivation-system* as it is. It is not sufficient to equip the I-Creature with a concept of time and with knowledge and ideas about a future, in particular about future needs that will become present needs Gust as the one that is experienced right now), if there remain determinate and unequivocal forces that compel it to engage in particular activities and actions all or most of the time.53

First we need, it seems, some mechanisms for dampening the forces of the system's original motivational mechanisms in order to make it capable - to some extent - to *resist* say current hunger. Present interests may retain the same kind but less forceful effects. Hunger, for instance, can still exist but not (primarily) as an irresistible force that directs behaviour but more as an inclination to deal with the interest in food and eating. (See Bischof, 1985, p.550.)

What effects shall *potential* interests then have in order to be evaluated as important and to be cared about? Shall a potential interest in food produce the same effects as a present interest in food, i.e., hunger? In principle it may function that way. In this case then, *just as* current tiredness causes the system to be concerned with thinking of how to get to sleep, how to find a place for sleeping, etc., imagined potential tiredness would similarly direct the system's thinking so that it ensures that it will have somewhere to sleep at a later moment of time.

But another possibility for motivating at least some kinds of anticipatory planning is to let representations of *potential primary* interests produce *present secondary motivations.* For instance, the system's imagination or conception of coming hunger or a potential threat can produce not hunger or fear but anxiety or curiosity, and an urge to find out about and think about this potential problem. The point is that anxiety and curiosity (as well as hunger and fear, etc.) are kinds of messages or formats for messages that the motivation-system *already deals with.* They are messages of what is to be cared about, and to act in relation to. And my point of departure is to try to make as much use as possible of the original motivation-system, with its mechanisms for information processing and evaluation. If there are motivational mechanisms on a more basic level than the level of reasoning, planning and reflecting (more built-in, more hardwired, phylogenetically earlier, etc.) it is an advantage (from an evolutionary point of view and from a designer's point of view) if such activities, anticipatory planning in particular, can be admitted and sustained by these mechanisms.

The secondary motivations in the original motivation-system indeed turn out to be suitable for motivating or driving anticipatory planning. The reason for this is that anticipatory planning - at least many forms of it - can be seen as *exploration, manipulation and play in an internal environment.* Anticipatory planning involves exploration of, manipulation of and playing around with representations of future and past events and experiences.

---

53 It might be relatively easy to fit in same anticipatory planning during the time that the 1-Creature would 'only rest' (and not think of anything), but this leaves the problem of subsequent plan execution or acting upon plans unsolved.

And, just as in (external) play and exploration, this activity does not occur primarily or only in relation to any present serious interest. It is rather an exploration of potential interests, problems and solutions. What we do (as constructers of an anticipatory planner) is to graft onto or exploit the motivations for exploration and play; boredom, restlessness and curiosity on the one hand and insecurity or anxiety on the other (which fundamentally relate to the interest in novelty and the interest in familiarity and security). (Cf. p. 27.).

Let us start with *insecurity.* Insecurity, due to lack of familiarity with a situation, object, area, etc., can (see pp.26-27) under certain circumstances motivate an individual to engage in exploration and manipulation in order to become more familiar with and achieve some control over the situation, object or area. Even when it experiences a situation as unfamiliar, novel and frightening, the individual may namely be motivated to explore the situation by a desire for excitement, a desire for challenge, a desire for taking risks etc.[54] In a similar way, the idea or conception of the *unknown and uncertain future* (cf. p.131), may motivate an individual to engage in manipulating and playing around with possibilities concerning his future. Planning and structuring the future, filling it with potential events and actions, producing ideas of what is going to happen, what one will do, etc., are ways of 'gaining control of one's future' and moderating insecurity.[55] Second, *boredom* and/or *curiosity,* which basically means that the organism does not experience sufficient stimulation in the form of new information and experiences, can in the anticipatory planner motivate imagination, planning and daydreaming as well as play and exploration. It can be interesting, stimulating and enjoyable to think about the future.

Most of the planning activities described on page 121-125 could be motivated in the individuals in these ways. These anticipatory planners engage in planning in situations where an immediate planner would be engaged in play, exploration or 'justresting'. Therefore the A-Creatures spend more time resting and engaging in planning and imagination but less time playing and exploring the external environment than the I-Creatures. In sum, I suggest that at least a main part of anticipatory planning is an internal correspondence to external exploration and play, and motivated in the same way.

Several of the *difficulties* that I spoke of in connection with immediate planning motivation reappear with anticipatory planning and in an amplified form. First there is the risk that an individual engages in *too much planning.* This may easily happen in an anticipatory planner if it is not carefully designed. The reason is that, in this case, there is no immediate need and motivation for acting that can become dominant and force the system to terminate the planning activity. In this way it is possible to dwell much longer in anticipatory planning, trying out many alternatives, trying to optimize the solution. The risk is increased as anticipatory planning often gives some satisfaction and can be enjoyable. One relatively simple remedy, however, that may counteract the risk of too much planning is that the planning may become tiresome for the system. Second, there is also an increased risk of producing inadequate plans due to the *longer time spans* that the anticipatory planner deals with. This increases the risk for the making of plans where the planner lacks

54This happens when the unknown is not *too* unknown and frightening. Otherwise the reaction may be to attempt to withdraw or flee.

55However, this motivational mechanism will be most relevant for sophisticated anticipatory planners in complex environments who are dealing with long term perspectives.

enough information. (lts anticipations are wrong, it makes plans for situations that are not predictable and not surveyable, etc.) In an anticipatory planner, it is therefore more essential than in an immediate planner to have an ability to also produce *sketchy* plans and leave it at this (avoid going down to the level of routines and preformed programs), and to make plans that deal with alternatives and leave various options open.

These two risks may be very *serious* in a more complex anticipatory planner than an A-Creature. They are not so critical though as long as we have an anticipatory planner with only a small number of interests to handle, where the environment is not very complex (as the environment of the Berry-Creatures that lacks certain complexity in that it is nota social environment) and where the environment and form of life of the creature sets limits in the size of time spans that can be relevant to consider.

Recall that my discussion of motivation of planning in an A-Creature is based upon the idea of maintaining and using the motivation-system of the 1-Creature as much as possible. And the motivation-system of the 1-Creature, in turn, the motivation-system of a *non-planner* that is retained. An alternative approach would assume that planning involves a more radical departure from original motivation mechanisms. Planning, one can argue, concems rationality. It means *thinking* and *reasoning* about what to do, when to do things and how to act, etc. And rationality, knowledge and *planning* make systems capable of leaving more primitive behaviour governing mechanisms as drives, impulses, feelings, emotions, etc., *behind.* They do not need such mechanisms to tell them what is important and when to do what. In brief, cognition will replace emotion.

I believe this wiev is mistaken. In my view, a planning capacity in biological systems is only a kind of overbuilding that *modi.fies* but does not replace (these) more fundamental or primitive mechanisms for behaviour control. Let me add a few words on this, by contrasting two different views on motivation. The first view is illustrated by the ideas of S. Schachter (1962). According to Schachter, *all* evaluations of what is going on, what is important, what is to be done, etc., are done by sophisticated 'high-level' interpretation. Emotions for instance, have some biological low-level aspect, but this is just *one* kind of *general* noise or bodily arousal. Toere *is* no fear, anxiety, curiosity, etc., without *rational* evaluations. It is by reasoning and explicit interpretation of the situation that individuals label the general noise as fear *or* anxiety *or* anger *or* happiness, etc. And thereafter they can act. Until one has reasoned and thought about the situation, there *is* no anger or fear, etc. In sum, there is some general physiological, low-level noise that is interpreted by high-level cognition. (See figure 3la.) Contrast this with the second view, namely, the view that on the lower 'biological levels' there are already *speci.fic* messages, and *specific* information, that may be handed over to high level cognition to deal with. (See figure 31b.) According to this alternative view, some valuable and distinct information is processed also in more primitive motivation mechanisms. Sometimes - as in pressing situations - this is drawn upon as such, and sometimes it is handed over to higher levels of evaluation.

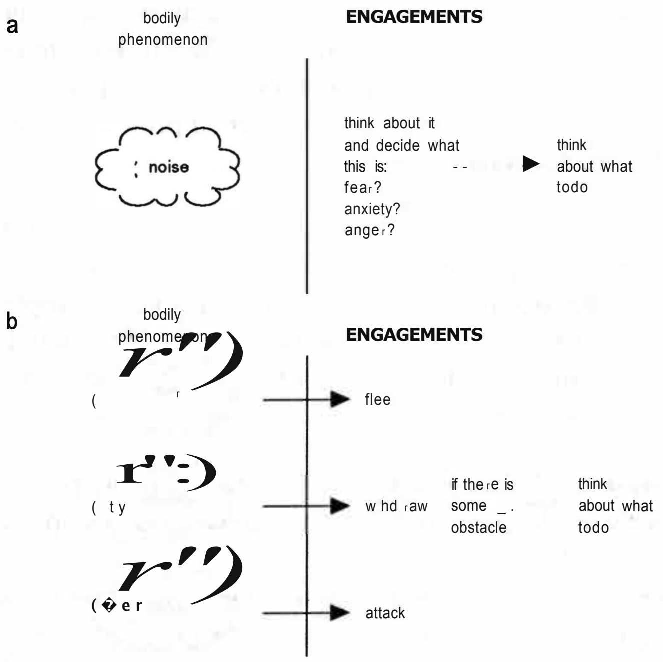


Figure 31. Two views on emotions as motivational mechanisms

From an evolutionary perspective the second alternative is the most plausible. First, we should expect ratiomorphous (see p.37) behaviour-regulating mechanisms that are older than mechanisms of explicit thinking and reasoning in individuals, and it indeed seems that there are such mechanisms. Second, given that some such ratiomorphous motivational mechanisms have developed during the course of evolution, these can hardly be expected to be left behind and be completely replaced because of the development of 'rational thinking' and high-level cognition. From a biological perspective it would be unwise to let an individual system be capable of completely building its own ideas about what is important and what to care about - and this is what motivation is about. This would be unwise regardless of the individual's capability for building a representation of the world and of itself and regardless of the powerful reasoning and planning capacities it may possess.

5. *A Qualitative Change: The Creature in Time*

To sum up the preceding sections, two steps are decisive for developing an I-Creature into an A-Creature. The first is to endow it with an *explicit time representation* and relevant knowledge organized by this. The second step is to *modify the motivation-system;* in particular, by giving up the one-primary-motivation-at-a-time in order to dampen present-time motivations. These changes are both required to ensure that potential interests can become commensurable with, and compete with, current interests, on a cognitive as well as on a more basic motivational level.

I further hold that this transition from an immediate to an anticipatory planner is a *qualitative* change. It is not just a *continuous* transition brought about by some quantitative extensions, such as endowing the system with a longer time perspective, increasing its memory, enabling it to consider a greater number of actions or means-end-relationships, etc. It is not a question of 'doing the same thing' but introducing quantitative changes to make it run more efficiently. On the contrary, it is a transition that demands both the introduction of a *new element,* namely, the explicit time representation, and a *reorganization,* namely, of the motivation-system where some basic principles of the motivation-system in an immediate planner must be left behind.

I do not claim that a qualitative difference will necessarily be *noticed* between *any* immediate and anticipatory pianoer. But the transition from an immediate to an anticipatory pianoer opens up for developments (without new qualitative changes) that can easily lead to what will also be *seen* as qualitative differences. It *opens up* for *new kinds* of phenomena and behaviour, as I shall try to illustrate.

First, the immediate planner always has one clear *overall goal* - getting. some food, fleeing from an enemy, etc. And planning basically means *'planning how'* - plan bow to realize this goal. The anticipatory planner, on the other hand, may be more concemed with *relating* several *disparate overall goals and activities* (considering bow and when to relate one activity or goal to other activities and goals). In this way it also engages in what could be labelled *'planning what'* - to plan what to do duringa certain period of time, or *until* some other thing happens or is to be done (as in example (A. l) and (A.5)), or plan what to do *next* (as in example (A.4)) - and *'planning when'* - as in example (A.6) where the individual thinks about *when* to build its night-shelter. Anticipatory planners are notjust able to *plan their actions* but are also, one could say, able to *'plan their time'.*

Second, and related to the first point, an anticipatory planner may have, or gain, some conceptions of *'waste of time'* and of *'saving time'.* For instance, a person may realize that before her friends arrive there is no time for both making <linner and getting those extra mattresses for the guests to sleep on tonight. Or an A-Creature may realize that before it gets dark it cannot both get material for mending its night-shelter, collect some food and work on the construction of the shelter. Correspondingly, the anticipatory planner may form conceptions like 'organizing my time' or 'being efficient'. For instance, 'if I do the things 'in the right order' I may arrive at accomplishing *more* and reaching more of my goals'. In this way the anticipatory planner may conceive of something as a waste of time or a saving of time. For instance, going to a distant food site when another is nearby may be recognized as a waste of time. And doing ***a*** before doing *b,* even though *b* may be more pressing or urgent than ***a*** but the time spent on both actions together will thereby be less, may be seen as gaining time, as in example (A.3).[56]

Third, the phenomenon of producing plans that one will never carry out can be given new dimensions with anticipatory as compared to immediate planning. The reason is that anticipatory planning means dealing with non-immediate problems, with hypothetical situations and with longer time spans, which creates the possibility for the construction of plans with the character of loose strategies and plans 'in case', etc. The anticipatory planner may indeed produce plans that it hopes it will never act upon, like a plan for what to do in case of fire, or in case I find a burglar in my summer-house, or in case it will be extremely cold for a period of time, etc. It may even produce plans that it *knows* it will

56 The basis of the conceptions of waste of time and saving time may relate to conceptions of wasting and saving *resources* like forces, efforts, etc.

never execute, nor see executed, like a plan for how one's funeral shall proceed. [57] In an anticipatory planner, you could say, the planning of action can be an end unto itself. Anticipatory planning must be motivationally more self-sufficient than immediate planning and be initiated, sustained and reinforced because it gives some form of satisfaction *as such,* that is, without being connected to any action or external behaviour (cf p.136). It can be enjoyable or fun to solve potential and hypothetical problems and engage in imagination, and there can be satisfaction in anticipatory planning because it reduces boredom and anxiety. Not only the actual reaching of a goal; but also a simulated, imagined reaching of a goal, it seems, can render satisfaction. Furthermore, this seems to relate to the following particularity of anticipatory planning as contrasted with immediate planning. Some anticipatory planning, at least, involves the capacity to *deliberatley postpone* the satisfaction of a current interest by *refraining* from acting upon a current motivation. An *immediate planner* may *wait* because the current situation is as it is, but anticipatory planing may also involve a *choice to wait* or to postpone. Now, what methods are there to motivate such *active* postponement or waiting? It seems that one method is precisely to have the anticipation of and imagination of the accomplishment of a goal intrinsically satisfying as such. In that case a system may well postpone action and attempts to immediately reach a goal for a while, because it is also enjoyable just *thinking* about reaching it.

Fourth, an anticipatory planner can be motivationally very unstable. Remember that in the A-Creatures, as compared to the I-Creatures, I introduced a dampening of the motivational forces for behaviour and gave up the principle of one-motivation-at-at-time. This was necessary to obtain a simultaneous consideration of *one* current - *the* current - and one or several potential motivations. But with this is also created the possibility for a simultaneous consideration of *more than one current motivation.* If there is a physiological need for both nutrition and for sleep, say, it is not necessary that hunger *or* tiredness will dominate and completely determine behaviour and attention for some time (as is the case in I-Creatures). Instead, both motivations can appear, but less forcefully and more as competing *inclinations* for what to deal with than as *determiners.* The anticipatory planner may be able to handle such a situation and not, as the immediate planner or the non-planner, vacillate for ever between the two choices; either remaining in the middle or rushing back and forth. The anticipatory planner may cope with, say, *simultaneous tiredness and hunger* by a *cognitive evaluation* by which it orders the interests and corresponding actions temporally. For instance, an individual may decide to get some *food first* and *then* find somewhere to sleep, or vice versa, depending upon what beliefs and ideas it has concerning hunger and tiredness. However, to leave the principle of one-motivation-at-a-time and to dampen motivational forces implies a general *decrease in motivational stability.* The number of motivations that can be entertained at one time is *increased* and their individual forces are reduced. Hence there are new possible forms of vacillation, hesitation, and so on. (Cf. Bischof, 1985, p.550.)

Furthermore, in a creature without a (the original) firm and biologically hardwired determination as to what goals to pursue and devote resources to, there is a breeding ground for new kinds of goals, more or less removed from the original motivational-experiential basis. Systems may indeed set up goals for which there is no such basis. To take some

---

57 This, however, is not a plan concerning ***one's own*** actions.

extreme examples, an individual might set up the explicit goal of proliferating his genes as much as possible, or set up the goal of dying (to plana suicide).[58]

The possible developments mentioned above are based upon the two basic properties that distinguish the anticipatory from the immediate planner; the changed motivational organization and, in particular, the *explicit time representation.* With sophisticated anticipatory planners with an elaborate time representation there is a breeding ground for creatures that are strongly influenced and marked by this time representation. Their perception of the *present situation,* on an experiential as well as a reflective/contemplative level, (and their actions in this present situation) may be strongly coloured by their own *representation of thefuture and the past.* On the *reflective leve/,* they may explicitly relate their present to their past and future; retrieve stored experiences from episodic memory and think and speculate about the future and interpret the present situation accordingly. On the *experiential leve/,* representations of their own fu ture and past may produce anxiety, anger, curiosity, satisfaction just as perceptions of a present situation can. (Cf. point 3 above.) If this property is pushed far enough, one may indeed get a paradoxical kind of biological creature, namely, a cre ture that can let its imagination and thinking dominate over action and let its thoughts about the future dominate over thoughts, perception and behaviour about and in the present (only worrying about tomorrow but never about today, 'enduring' the present because it 'looks forward' toa hetter future, etc.) A creature that is thus heavily determined by its time representation, on a reflective/contemplative as well as on an experiential_ level, I shall call *a creature in time.*

As a last point, yet another phenomenon that can appear in an anticipatory planner is that of *training.* Both immediate and anticipatory planners engage in play and exploration. An immediate planner does so due to a current interest in new information to deal with, and it is driven bya present motivation, boredom, curiosity, etc. But the anticipatory planner- and only the anticipatory planner - may also decide to engage in exploration, or in climbing trees, say, in order to become more skilled or learn something that it may use later. It may engage in training fora particular purpose. (See example (A.5).)

With this I conclude this chapter and the attempt to present a system explanation of anticipatory planning. In the next part, I deal with the evolutionary explanation of planning, immediate as well as anticipatory, discussing the roots and the evolutionary value of planning. Finally there isa chapter devoted to the *human* planning of action.

58The fourth point also ties into the first point. With anticipatory planning one is likely to find planning that involves decisions about *what* problem to solve, *what* goal to pursue; in other words, planning that is in the first place a 'planning of what' and nota 'planning of how'.

PART THREE: AN EVOLUTIONARY UNDERSTANDING OF PLANNING

['llie liorse}decided to p[an out a nice ref d day for itseif. 9L fittfe trot fater on, it tliouglit, may6e arountf tlireeisli. .91.fter tliat a 6it of a fie down over on tlie east side of tlie fie[d wliere tlie grass was tliick!,r[..}It afso quite fiÆr,d tlie nation of spending lialf an liour wafkJ,ng afternatefy a fitt{e 6it to tlie [eft and tlien a fittfe 6it to tlie riglit, for no apparent reason. It didn't k,now wlietlier tlie time 6etween two and tliree wouU 6e 6est spent swisliing its tau or murang tliings over. Of course, it couU a[ways M 6otli, if it so wislietl andgo for its trot a fittfe fater[...}

(jood. 9ln e ceffent p[an. .91.nd tlie 6est tliing a6out it was tliat liaving made it tlie liorse couU now compfetefy and ut-terfy ignore it. It went insteadfor a [eisure(y stand under tlie on{y tree in tlie faU.

'IJougfas .91.aams, 'IJirf((jent{y S lio[istic cktective agency

10. ON THE EVOLUTIONARY ROOTS OF PLANNING

1. Introduction

A capacity to plan, as I conceive of it, is an advanced form of behaviour control. As such, it is not a phenomenon that emerges out of nothing with its own new mechanisms. It does not make its entrance in evolutionary history as 'a whole new package' but builds upon mechanisms and capacities that were already available. In this chapter I discuss some of the mechanisms and capacities that may form a background for the emergence of systems that - at times - make plans for their actions. In this way, I hope to gain a greater understanding of the planning of action as a biological phenomenon. I shall discuss three 'stages' of behaviour control, where some new control mechanisms are added in each stage. I call these *the instinct system, the trial-and-error system* and *the playing system.* [1]

As in previous chapters, I make use of the example of the fictitious Berry-Creatures. But I believe that most of what I say is biologically relevant and plausible, even though, of course, the three stages, like all stages, are only theoretical tools and not something one can go out and look for in nature. It is also important to be clear about the following: I am not claiming that the stages that I describe, with their respective mechanisms and capacities, are *necessary* preconditions for the development of a planning system. I only argue that given that they were available in some species - as I believe they were - they were useful in the evolutionary emergance of planning.

We are looking - at least - for the following competences:

(1) An ability to represent a problem- a discrepance between start and goal- and devote resources to it by thinking about it and trying to represent a way for salving the problem (and, for anticipatory planning, to do so also for *potential* interests and problems) which requires:

- an ability to *represent possible situations* - in particular *goal* and *start situations*

- an ability to *represent possible events,* in particular one's own possible *actions:* to represent their preconditions and consequences or, in other words, which situations they can. transform and how (which requires knowledge of the spatial structure of one's environment and of regularities in one's interactions with the environment)

- an ability to *predict* events - that is, to relate preconditions to outcomes, to foresee consequences of events and actions

- for anticipatory planning, an ability to 'construct' problems in the sense of predicting and representing problems *not* related to a current interest and a corresponding motivation for acting

[1] Compare this chapter with Lorenz' discussion of *the roots of conceptual thinking* (Lorenz, 1973, pp.158-211.) to which this chapter owes a great deal.

(2) An ability to generate representations of actions and goals without immediately acting upon them

(3) An ability to deal with representations of *alternative* courses of actions (which are all related to one and the same goal) and to *evaluate* and *compare* such representations

(4) An ability to conceive of means-end-relationships (or subgoal-goal-relationships) and to relate two or more actions to one another by such relations

(5) An ability to compose representations of behaviour pattems by decomposing other such representations and using their parts

2. *The First Stage: The 'Instinct System'*

The 'basic system' or the 'instinct system' is a relatively primitive kind of system with limited behavioural flexibility. For each of its biological interests, it has one or several *instincts.* An instinct, recall, (p.23), consists of three parts or mechanisms: a motivation in the form of a *drive* (that relates to the interest), *afixed motorpattern* anda *detector* for detecting the relevant situtation where the pattem is to be applied. That is, these motor patterns are accessible only in *certain,* relatively well-defined, situations.

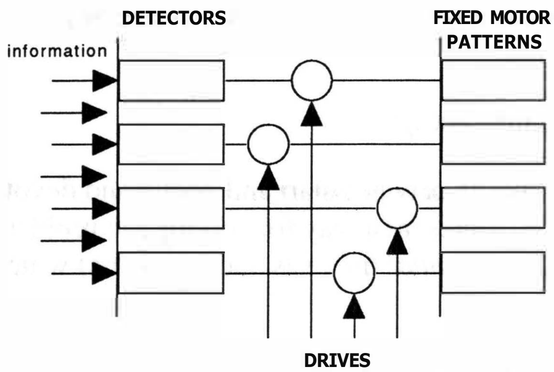


Figure 1. An instinct

The success of instincts as mechanisms for behaviour control is due to the existence of certain permanent or long-term regularities in the environment. Consider as an example the mosquito that bites everything that has a temperature of 37 Co and smells of butyric acid. In the mosquitos' natural environment these properties regularly coincide with the property of being a suitable source of nutrition for the mosquito. The function of a *detector* is to detect situations that, with certain *probability,* belong to a biologically adequate kind of situation, without incurring too much information processing. However, there is knowledge that cannot be supplied in this way, and there are certain aspects of an organism's behaviour that cannot be determined in this way. Even an instinct system must have some supplementary mechanism for behaviour control. First of all, there are spatial regularities conceming location and how sites are related to one another in the system's particular environment. Even though there is no great variability in *how* an instinct system moves, *where* it moves is an issue that must be determined in some way or other. It is also evident that it will hardly be the case that whenever a particular drive arises, the relevant situation will just 'be around' to be detected. The *fixed* motor pattems of the instinct system - which are often some *consummatory* or end behaviour - must be complemented with some kind of *appetence* behaviour for *moving* and getting to a situation relevant to a drive (see section 3.2). Some behaviour elements for - simple - *locomotion* must be more generally accessible to the system. It is also clear that as soon as organisms are able to move around, it is of selective advantage to have good means for spatial orientation. Although there are some fundamental movement or orientententational mechanisms such as taxis, kinesis, etc., there will be evolutionary pressure towards more advanced means for orientation that exploit regularities in the spatial environment; i.e., for mechanisms that allow an individual system to orient itself in its environment on the basis of its own experiences. Second, there is knowledge related to regularities in the environment that are so 'new' that nature cannot have considered them. For instance, a new kind of enemy may have appeared in the environment. And third, there is knowledge of regularities that *is not* 'built in', even if it 'could have been', for reasons of cognitive economy.

In order to have systems utilize regularities of these kinds, where the corresponding knowledge is not provided for genetically, we must introduce some kind of *learning.*[2] Furthermore, this learning must be of a simple kind. Toere must be room for it in a system that is only capable of stereotype consummatory behaviour and of appetence behaviour limited to 'simple locomotion'.

Let us assume that we want a basic system to make use of the following regularities in directing its behaviour:

• Food that it eats and recognizes because of its form and colour also has a characteristic smell (this could be useful for finding food in darkness);

• When it perceives a dead conspecific, it also perceives a certain acrid smell (this could be useful forevading <langer);

• When it perceives the straightness of a horizontal line in different parts of the visual field it shortly thereafter perceives a dangerous cliff;

• When it perceives a particular shape approaching, it shortly thereafter perceives an enemy.

We want the system to extend the range of situations where it applies its instinctual behaviour patterns: to apply its eating routine on things with the mentioned characteristic smell, react to the acrid smell as to the perception of a dead conspecific, apply the same action pattern when it perceives the straight lines as when it perceives a cliff, and apply the flight routine when it perceives the particular shape.

---

2 Where learning means that *an individual's experiences* cause an adaptive change in the probability of the occurence of a particular kind of behaviour in a particular kind of situation.

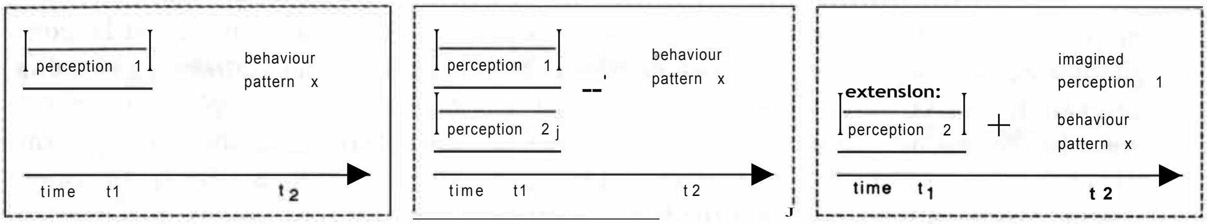


Figure 2a. Association of simultaneously occuring perceptions

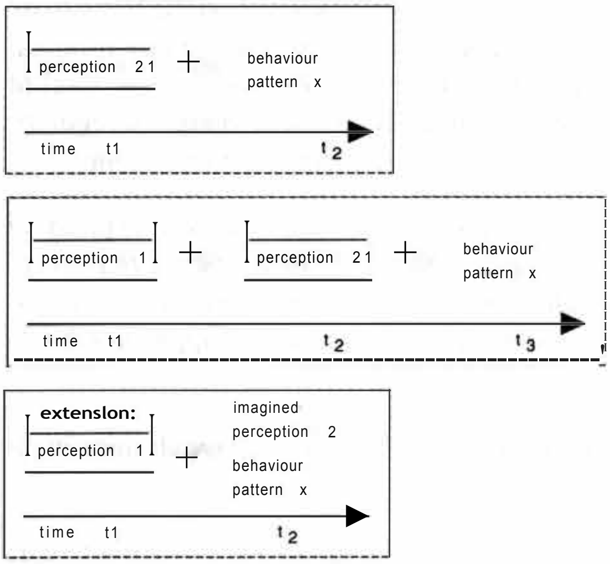


Figure 2b. Association of following perceptions

In other words, we want some kind of *associative learning.* Two such kinds of leaming are relevant. In the first two examples we want association of *simultaneously* occuring perceptions (see figure 2a) and in the latter two, association of closely *following* perceptions (see figure 2b). It is not too difficult to imagine some mechanism that could deal with these kinds of associative leaming. For the first kind, one can connect all the perceptual sub-systems with some connections where there is normally no activation spread because of 'high resistance'. The resistance can be assumed to remain high *unless* the two systems or components that it connects are simultaneously active. In this case, the resistance will decrease. Now, some perceptual sub-systems - say, the one responsible for perception of the colour green, for the form of the food, or for the particular smell of the food - will often be activated together. Because of the mechanism described, every time the individual perceives the particular form and colour, it will then also imagine the smell, and whenever it perceives the smell, it will imagine also the form and colour - and this will make the eating behaviour available. Similarly, when it sees the particular straight lines, it *will.filfin the rest, imagine the cliff-* and apply its fixed motor pattem.

For the second kind (association of closely following perceptions), again connect all perceptual sub-systems but now with double connections of a kind that conducts acitivty in only one direction. Here we can let conduction increase in connections between such elements that are *active successively within a brief time.* And so, when the individual perceives the particular shape approaching, immediately thereafter, it imagines the enemy and flees, and so on. Toere is an activation of internal elements corresponding to regularly recurring sequences of events - *but at a more rapidpace.* Thus we obtain a basic form of*prediction* in the system. (See figure 3.)

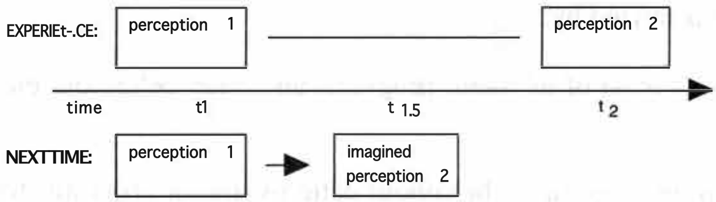


Figure 3. Prediction

As one may realize, these learning mechanisms can also be used in leaming spatial regularities. If perceptions of particular landmarks regularly occur together with perceptions of certain things, events or activities, one may associate them; and so, for instance, when one perceives a particular landmark, one also imagines the food that usually is there, even if, say, some grass is hiding it. And if the perceptions of two landmarks regularly occur in succession when one walks a certain path, one may imagine *the next* landmark upon only seeing the first one. In this way, a kind of 'small cognitive map', a *local cognitive map,* useful for orientation and for locomotion, may be established.[3]

The mechanisms described handle the learning of regularities that are frequently experienced or perceived. But with some additional mechanism for evaluating the importance of experiences, it is possible to let rare but important experiences influence connectivity between nodes as well, and in this way also enable a system to learn from a single experience.

In sum, even a relatively *simple* kind of system as the one described here, can have much goal-directedness in its behaviour, some due to its genetical endowment but some also caused by an individual system's own experiences. Its appetence behaviour does not have to be limited to a random moving around. It may make use of a limited capacity for prediction, and of some kind of local maps. Already in such a system there is a basis for some of the requirements fora planning capability. It has an ability to re-present certain perceptions. It has knowledge of - and can represent- certain regularities and can leam about such. And it has a limited ability for prediction, which it uses for behaviour control.

Comparing this discusson of an instinct system with the discussion of perceivers-and-behavers of different orders (section 6.7), *instinct systems* can be any system up to the level of routine-govemed systems.

3. The Second Stage: The 'Trial-and-Error System'

In a *trial-and-error system,* compared to an instinct system, there are many more aspects of behaviour that are *not* genetically determined, and there is less determinacy as to *when*

---

[3] The kind of perceptual tilling in of information that I have described are examples of what I called *closure* in section 7.3, where I spoke of this as an element for a planning competence.

a certain behaviour pattern can and will be applied. Such a system has a capacity for *trying various* behaviour patterns in certain situations, even new behaviour patterns.

The trial-and-error system is characterized by:

(1) A more general accessibility of most of its motor programs and other behaviour elements

(2) A capability for using *parts* of more complex behaviour patterns (motor programs for instance) *separately* and *combining* such smaller behaviour elements into *new* complex ones

Compared to an instinct system, the *motor behaviour* of a trial-and-error system is more differentiated and flexible. In an instinct system, remember, it is only simple movement patterns for locomotion that are generally accessible to the system, whereas all other aspects of behaviour - more or less complex stereotype motor sequences - are restricted to well-defined situations. The basic advantage of a trial-and-error system over an instinct system lies in its potential for dealing with situations for which nature has not provided any instinctual solution and for which there is thus *no suitable motion pattern.* It may deal with *obstacles* to the performance of hardwired motor patterns and, in general, with unusual or divergent situations which nature has not foreseen. If a trial-and-error system is hungry and finds something that is similar in form, smell, etc., to some familiar food but enclosed in a hard shell, it may start biting it, scratching it, hitting it with a stone, throwing it on hard ground, etc., and might thus succeed in cracking it. Or if, say, it gets enclosed, by a human being, in a box, it might do various things with its feet, mouth and limbs, like scratching the walls, putting its feet, hands and nose in the fissures and may be even succeed in opening the box and getting out. Yet the trial-and-error behaviour occurs within the frames of instinctual behaviour. The motivational element (the drive) must be there, and often the trial-and-error behaviour ends with some consummatory behaviour.

The examples above illustrate relatively sophisticated trial-and-error behaviour, but there are also more primitive instances. Think of a system that is located behind some obstacle and wildly applies *any* behaviour patterns it has accessible, like closing its eyes, licking its feet or kicking a stone away. Indeed, if one conceives broadly of trial-and-error, all *appetence behaviour* in instinct systems is a very basic kind of trial-and-error-behaviour. (Think of a butterfly flying against a closed vindow, randomly changing its location.)

A trial-and-error system will be more adaptive if novel behaviour patterns that are successful are incorporated into the system's behavioural repertoire for the appetence-phase concerning the interest and problem in question.[4] In other words, it will be more adaptive if it *learns* from *successful trials.* Learning in a trial-and-error system can be of a more complex kind than the learning achieved in an instinct system. An instinct system may start to use some fixed, 'habitual' behaviour patterns in some different situations than it has done before. But in a trial-and-error system a *new* behaviour pattern may be associated to some kind of situation (often novel or somehow unusual to the system).

[4] But note that trial-and-error behaviour is possible in principle without being related to learning, and it can be advantageous as such to the system. However, the adaptiveness will ***increase*** with learning.

This opens up some important possibilities. When one is no longer restricted to the application of *preexisting organized whole* behaviour pattems but also *composes new patterns* by combining parts and leams such new compositions, there will be a *value* in ways for relating various *parts* of a sequence to each other. It is not just interesting to associate a new situation to a habitual behaviour pattem and a start situation, but to relate new situations to behaviour elements and to an end goal. To do so, behaviour elements may be related as means to end or subgoal to goal. It will be of value to recognize 'situations on the way', which are outcomes of *apart* of a behaviour pattem or sequence. (Cf. the discussion about trial-and-error behaviour in section 6.7.) The point is not that a trial-and-error system *necessarily* develops these capabilities, but that trial-a1;1d-error allow for their possibility by giving them a function.

In which ways does this stage then give an extended basis for a planning capacity? First, if planning means to intemally *try out* various action pattems, it would also seem to be the case that there must be a corresponding capability of *trying out* actions *in behaviour,* which in tum requires a *repertoire* of *available* behaviour patterns - motor patterns. And so, an important basis for planning in the trial-and-error system lies in the increased behavioural repertoire, in the accessibility of behaviour and motor elements and in the capability to combine behaviour elements into new pattems.

Second, trial-and-error involves the ability of trying *several alternative* behaviour patterns for reaching one goal situation and possibly an ability to leam from such trials. The increased potential for leaming various regularities, including such that involve *new action patterns,* is an important basis fora planning system. This is also true for the created possibilities of recognizing means-end-relations or subgoal-goal-relations.

Third, in general, the more differentiated and refl.ned motor capabilities in the trial-and-error system - as compared to the instinct system - also imply a pressure for *a more precise and detailed representation of the environment.*[5] This, as well, is useful fora planning system.

It ought to be clear that there can be trial-and-error systems of widely varying degrees of sophistication and flexibility. Indeed most animals above insects and 'upwards', to cats, dogs, and so on, could probably be characterized as trial-and-error systems. What differs between species are factors such as the extent to which they are capable of taking established pattem parts and composing new patterns and how powerful their learning capability is. And there are differences regarding the extent to which the behaviour of a species is of the instinct type and bow much is of the trial-and-error type.

---

[5]Indeed ***both*** the evolution of ***nwtor abilities*** and the evolution of the ***representation of the spatial environment*** are adaptions to demands from a life in a complexly structured spatial environment, and they evolve hand in hand. (The ***less homogenous*** the spatial environment is, the greater the need is for smaller separate behaviour elements for building larger pattems adapted to particular situ tions, and the nead is also greater for more sophisticated representations of the environment.) (See Lorenz, 1973, p.171 and p.176.)

4. The Third Stage: The 'Playing System'

The third stage on the evolutionary path to planning is a trial-and-error system with a capability of *playing*. I have mentioned playing in previous chapters, in particular in the chapter on biological functionalism where I described it as a biological phenomenon, and in the chapter on anticipatory planning where I discussed parallels between playing and anticipatory planning. I will now, however, provide a more detailed description of the phenomenon and discuss bow it can contribute to a basis for a planning capacity. In my view, the development of a capacity for playing *considerably* increases the strength of a behavioural system and in several ways implies a move towards systems capable of planning.

Let me start with some examples of play behaviour in the Berry-Creatures.

(1) (playful spatial exploration)

A Berry-Creature moves around, goes to various sites, looks at things, sniffs at objects, creeps into cavities, goes around trees, climbs trees, jumps into trees, etc. - and seems to be up to notbing particular. Here we have an *exploring* individual who is not looking for anything particular but just looking around in order to familiarize itself with the environment; to see what is around and where it is.


Illustration 10. Playing

(2) (playful manipulation)

A Berry-Creature finds a big coco-nut which is something that it has never seen before. It rolls the nut, kicks it, rubs it, throws it, hits it with a stone, etc. It arrives at cracking it, and then tastes it - but does not eat it.

(3) (playfully dealing with a potential obstacle)

A big tree has fallen down over the path that a Berry-Creature is walking on. It starts to climb over the trunk. When it reaches the other side, it climbs back, then starts climbing over again, now a bit further down the trunk. It repeats this at different parts of the trunk. It is difficult at some parts. Then it pulls and bites off a branch, throws it over the trunk, repeats this, then starts climbing over again, and so on.

A defining characteristic of play or playful behaviour, in my view, is that it is *not serious.* Play behaviour is *engaged in* irrespectively of any primary or serious interest. Play - moving around, exploring, manipulating, etc., but with no particular goal or problem in mind - is driven by *secondary motivations:* boredom anda desire to see and do something new - curiosity - or a desire for challenge, adventure or excitement. (Cf p.27.) (Note, I here use 'goal' in the sense of something that is *perceived* or *conceived of* or somehow *detected bya system* and made use of by the system for *initiating* or *terminating* some behaviour pattern.) That these are *secondary motivations* means that an individual *only* engages in playing when it has no *serious* or *primary* problems to deal with (cf. p.26). Yet many instances of play can be characterized as *'non-serious problem solving'or* as 'non-serious problem solving with *loosely determined goals'(in* contrast to planning which is *serious,* goal-directed problem solving). In the examples above, the goals might be 'see what is over there', 'to hide behind a tree', 'toget something rolling', 'to get over an obstacle', 'to throw a branch over a trunk', etc. There are two senses in which these goals are 'loose' or non-serious:

(1) The *play activity as such* is not *initiated* by any conception or representation of these goals and corresponding problems. Rather these are problems and goals that come and go *during play.* They emerge out of the system' s current interactions with the environment, and they easily replace one another. They are more or less 'temporary' goals, which during a brief time, loosely direct the system's activity. One goal - for instance, rolling away a stone -can swiftly be replaced by another-like throwing the stone into a hole or jumping into a tree. A problem may be abandoned even if it is 'unsolved' because the system gets bored or because it happens to find something that is more interesting. When manipulating an object, for instance, no obvious problems or goals may be involved at all. It is more a question of simply doing *something* with the object. Furthermore, in the third example above, there is no problem or goal in a strict sense. The system does not *stay* on the other side once it climbs over the trunk, terminating 'the climbing over' and initiating some new activity, but it goes back to the first side and repeats the climbing. It continues to do the same thing - until it gets tired or bored.

(2) Not only are these 'loose' goals easily replaced by other 'loose' goals, but as soon as a *serious* problem turns up, they are immediatley abandoned or left behind_ (For instance, if in the examples above, an enemy turns up, or if it gets hungry or cold, the Berry-Creature will stop playing.)

In what way is playing important as a basis for planning? Playing implies an increase in knowledge, in particular, knowledge of such regularities that involve the system's own actions. Compared toa trial-and-error system which does *not* engage in any *play/ul exploration or manipulation* the playing system's *potentialfor learning* is considerably increased.[6] In exploring and manipulating sites and things, for instance, biting, tasting, creeping into holes, etc., the organism is not after *eating* or *hiding,* etc., but it learns about *possibilities.* In Gibsonian terminology it finds out about the *affordances* of objects, sites and situations, i.e., whether something is edible or not, where it is *possible to hide,* whether it is *possible to pass over* a trunk in such and such away etc. (Gibson,

6Notice that for a capacity for playing - in contrast to a capacity for trial-and-error behaviour - to imply a selective advantage, it *has* to be combined with learning.

1979.) In this manner it can gain competences and knowledge that can be of use *when* a serious problem situation arises, reducing the need for - possibly riskful - trial-and-error behaviour then. A playing system makes extensive use of its capability of trying *alternative ways* for reaching a particular situation, and possibly comparing them, like in the example above .where it tries various ways of getting over the trunk. With this follows possibilities of learning about goal-subgoal-relationships.

It should be noted that playing constitutes a particularly strong basis for *anticipatory* planning. Both are activities that do not relate to any current and serious interests but involve moments of *constructing a problem* not related to a present interest and a corresponding motivation. In the chapter on anticipatory planning, I have already indicated bow playing and anticipatory planning run parallel to each other in several respects (cf. pp.135-136). The motivational mechanisms for devoting time and resources to 'constructed' or 'imagined' problems (that is, what sets off, sustains and reinforces the activities) may, to a great extent, be the same for play and for anticipatory planning - curiosity, excitement, obtaining control over the unknown and unfamiliar, and so on. In particular, it is evident that a capacity for play forms a background for the kind of planning that approaches 'day-dreaming'. That kind of planning may indeed be regarded as *internalized play;* non-serious and loosely goal-directed.

Anticipatory planning, daydreaming and playing are all *anticipatory activities.* They are not immediately valuable in that they improve a current situation or solve an immediate serious problem, but they may 'pay off' in later situations. They involve exploration of and learning about potentialities or principles. One may learn specific things, like how to handle specific potential problems that may turn up in the future, or one may improve one's competences on a more general level. Playing often means learning motoric patterns and becoming skilled. And 'daydream-planning' may parallel this in improving mental skills (imagination, representing, reasoning, using one's memory, etc.)

Note that I am not just indicating abstract parallels or parallels of classification or description between play and anticipatory planning. The point is that certain already existing *mechanisms* and ways of organizing and making possible play behaviour in a system can be used if one wants to make an anticipatory planner out of the system.

As I pointed out in section 3.3, page 26, playing - as it occurs in the biological realm - is often *risk/ul.* While playing, organisms can get burt; they are easily exposed to 'enemies', pay less attention to <langers, etc. This seems to imply that a motivation and a capacity for playing as a compensation must endow organisms with a considerable selective advantage. The explanation is, I believe, that it does so because it is an outstanding method for the system to extend its knowledge and competences of different kinds.

Play is one of nature's great inventions. Here we see systems that make use of the principle of *simulating,* in the sense that they *reproduce* processes, events and relations between them and *experiment* with them, with some possibility of *speeding up* performance and without the *risks* that would be involved if one actually or *seriously* made those things happen. In this way, playing is a predecessor of planning. Planning systems are a later invention of nature where the principle of simulation is made use of even more forcefully. I am not proposing that it is impossible to build a planning system that is also an autonomous *agent* but which lacks a capacity for playing. But I *do* believe that this capacity is one of the clues to the sophisticated nature of planners and agents in nature, and consequently, it is not a capacity that should be deemed as a side issue. If we can produce an artificial system that plays in some interesting sense, we will be on the way towards producing an autonomous agent and a planner as well.

5. Conclusion

To sum up, I have, in three stages, discussed the evolutionary background for a planning capacity. I have *not* argued that these stages are *necessary* requirements, but rather that they may supply some basis for a planning capacity. According to the analysis above, we have at earlier stages the following:

In the *instinct system:*

- Knowledge of regularities and a (limited) ability to extend this knowledge
- Representations of perceptions (closure, associative memory, etc.)
- An ability to predict, and a mechanism for using this ability in behaviour control
- Goal-directedness of behaviour

In the *trial-and-error system* one may also find:

- Increased knowledge of regularities
- An ability to construct and learn *new behaviour patterns*
- An ability to decompose behaviour patterns and combine behaviour elements into new complex patterns
- An increase in the behavioural repertoire
- A mechanism for trying out several alternatives to reach a goal situation
- Knowledge of goal-subgoal-relations

Furthermore, *in the playing system:*

- Still more increase in knowledge of regularities and in behavioural competences
- A considerable increase in the behavioural repertoire
- An ability to construct problems and goals that are not related to discrepances currently detected by the motivational system
- An ability to devote resources to such problems

Note that in this discussion of the roots of planning I neither speak of linguistic nor social capacities. The reason is that I neither think planning is intrinsically a social phenomenon nor that it depends upon communicative capacities. For my purpose, conceptions of planning that presuppose such factors are not adequate (cf. p.40). I do not assume that the planning of action *in principle* requires these capacities, nor that they actually did precede a basic planning capacity in in nature. The basic phenomenon of individuals structuring their own future actions - thinking about what to do and how to do it - is, I believe, more deeply rooted.

11. ON THE EV0LUTI0NARY V ALUE OF PLANNING

1. Introduction

In the previous chapter I discussed three stages of behaviour control - the instinct system, the trial-and-error system and the playing system-which I believe make up a background for the evolution of a planning system. The search for the roots or precursors of planning is one part of the approach of biological functionalism. But the second central question of this approach remains: What is the *evolutionary value* of planning?

In this chapter I do the following: I first present a general discussion regarding the evolutionary value of planning, analysing constraints and requirements in general tenns. In the next section I discuss immediate planning. I consider some examples of the I-creatures' planning and point to the kinds of benefits or advantages that their planning renders them. In the fourth section I do the same for anticipatory planning, in the context of the A-Creatures. In the last section, finally, I discuss 'nonadaptive' planning of action.

2. General Requirements for an Evolutionally Valuable Planning Competence

The capacity of planning of action, as I conceive of it, emerged relatively *late* in the course of evolution, and did so in organisms that already had many other capabilities and strategies available for biological problem solving. And so the question is: What is the selective advantage of planning as a strategy for problem solving in systems that are already endowed with such other mechanisms, capacities and strategies? I have in mind the mechanisms and capacities in what I have called instinct systems, trial-and-error systems and playing systems respectively. (Or, to use some exampels from K. Lorenz' list of - increasingly 'open' - behaviour control mechanisms: (1) kinesis, (2) phobis and taxis, (3) imprinting, (4) sensitisation, (5) habituation, (6) trauma, (7) instinct, (8) unconditional reflex, (9) conditional reflex, (10) associaton, (11) motor leaming, (12) curiosity...) (Lorenz, 1973, pp.65-194.)

It is *not* the case that the *general* ability to 'think about what to do and how to do it before actually <loing it' or 'to structure ones activities beforehand' has evolutionary value. There are many *constraints* on a planning capacity for which this holds. Of course these constraints take a particular form for each single system, but they may also be classified or subsumed under more general descriptions.

An evolutionally valuable *planning capacity* must, first of all, be *realised* in *planning activity.* And this planning activity (in turn) must have some *effects on behaviour.* Endowing a system with the capacity to plan must, in a critical amount of cases or situations, lead to actions that *better* solve the syste.m's problems than if there were no planning involved in the action production. The planning capacity must lead to hetter problem solving in the sense that, thanks to planning, the system will obtain goals that it would otherwise not have obtained, or that it will obtain goals in a hetter way. In general, adaptive planning implies one or several of the following *gains:*

• A gain in time

• A gain in effort

• A gain in some (other) kinds of resources that are scarce or valuable

• A gain in safety

In the chapter on planning and problem solving in *The Handbook of Artificial Intelligence,* (1982, pp.515-516.), three points are listed as the general benefits of planning:

(1) The *reduction ofsearch.* For instance, if you plan, you may not have to go to the library twice if you borrow a book and retum one at the same time.

(2) The *resolving ofgoal con.flicts.* For instance, if you want to build a house you shold plan to put in electical wiring first and install the dry walls afterwards and not first install the walls, as this will preclude the obtaining of the other subgoal (of having electricity in your house).

(3) The *provision ofa basisforerror recovery.* A plan, it is said "can be used to monitor progress during problem solving and to catch errors before they do too much harm." *(Ibid.,* p.516.). If the agent does not find the state of the world as it had expected it may stop and maybe engage in replanning.

In my view, these three benefits do not really have the same status. It is true that some *reduction ofsearch* is always involved when a planner acts according to an adequate plan. Because planning involves deciding on *one* course of action, and *not behaviourally searching* for a solution, there is reduction of search. Toere is *internal* search instead of *external* search, and this may imply a gain of time, effort or security. In addition, because there can be a reduction of behavioural search due to the particular course of action chosen (if this is one that leads to a gain in time or effort over other altematives, as illustrated by the library example above), reduction of search will be achieved. The *resolution of goal conflicts,* as well, may be said to be a general benefit of planning, at least if one construes 'goal conflict' in a broad sense, not only as conflicts between independent overall goals but also as conflicts between subgoals. But the third benefit is, in my opinion, of a different kind. It has to do with a system's flexibility when following a plan and is not a general benefit of planners over non-planners, but of some planners over other planners. Think of a rigid plan-follower. Furthermore, non-planners as well may have an ability to recover errors.

Further constraints on a valuable planning competence in a system are of course given by the properties and characteristics of the system's *environment.* And in a *particular planning situation* it is given by the characteristics of that *particular* environment or situation. The demands on a valuable planning capability in a system are higher if one or several of the following is true of the system and its environment:

• Many relevant changes and events are not predictable, or are at least difficult to predict, for the system.

• The environment is unstable. Relevant changes happen *quickly.*

• Relevant aspects are uncontrollable. The number of events that are uncontrollable for the planner is too large. The planner has limited resources for control.

• The environment demands that the organism is almost constantly involved in a 'struggle for life' and is in the mode for perceiving-and-behaving.

• The environment is too simple and predictable (and does not provide many alternative possibilities for solving the systems problems). Here one will do better by using fixed solutions; habits, routines, and so on.

Note the following trade-off. The planning of action will be most valuable in a situation or environment that is neither *too predictable* (then one will do better with fixed solutions) nor *too unpredictable* (in which case the system will do better to flow with the tide, take opportunities, only concem itself with taking onestep at a time).

On the other hand, if a system lives in an environment that *is* predictable in important aspects; of which it has control; where some relevant changes *do* take time; in which the system does not have to constantly struggle for life, and which is rich enough in providing possibilities for problem solving, then the probability for the emergence of a planning capability is high - and the demands on the planning capability low.

3. On the Evolutionary Value of Immediate Planning

I will now in two sections speak of the evolutionary value of immediate and anticipatory planning respectively. That a planning capacity is *valuable* from an *evolutionary perspective,* means that this capacity, in a sufficient number of cases, renders individuals that are endowed with it a selective advantage (cf. p.19). But let me point out that there are other perspectives from which it may be interesting to consider the value of planning as well. One may speak about the value of a planning capacity, or of a particular instance of planning, from a *particular individual's perspective.* And here I want to distinghuish between *contemplative/rejlective* and *experiential* value. Take as an example the case of an individual's planning what to do the rest of the day. Why does the individual do that? Contemplative value is illustrated by answers (which need not be expressed of course) such as 'I have so much to do today, if I construct a plan I will be able to fulfill more of my goals' or 'By planning I will gain time'. Experiential value is illustrated by answers such as 'I *enjoy* structuring my time' or 'I must have some idea of what is going to happen, *otherwise* I get anxious'. One may also speak about the value from a *social perspective.* From this perspective, planning may be valuable because it makes it possible to coordinate actions and collaborate and thereby reach collective (as well as possibly individual) goals, or because planning ones action may make one think about what other individuals want, or becuase planning may lead to group cohesion, and so on. A given instance of planning can of course be valuable from one of these perspectives but not from another. But note that it is not the case that any of the perspectives mentioned are necessarily exclusive.

I will, in this and in the next section, focus on value from an evolutionary perspective. But in the last section of this chapter and when I speak of the value of human planning in chapter 12, I will also pay some attention to the other perspectives conceming social and experiential value.

The first exampel of planning on page 92 in chapter 7 describes an 1-Creature who has found a Red-Berry-tree with Red-Berries that grow on thin branches unusually high up and makes a plan for how to deal with this. Compare this situation with the performance of a trial-and-error system that is limited to behavioural trying and, if it engages in object manipulating of some kind, to the employment of material that is available in its immediate surrounding. It is easy to see that the strategy of the planning 1-Creature does not only imply a gain of effort and security (a trial-and-error system might climb up and down and throw stones or might try to get out on the thin branch and fall down), but probably also means that the 1-Creature, in contrast to the trial-and-error system, will solve this particular problem (obtain these Red-Berries) *at all* (although the trial-and-error system can of course leave the site and go and search for food elsewhere).

The second example describes an 1-Creature planning a route to getto a cave. It plans to follow a novel path which is a short cut. This example primarily illustrates the benefit of time gain, with a reduced risk that the individual will freeze to death. (Compare this with a system of Berry-Creature kind minus a planning capacity which would in this situation follow 'the usual path'.)

In the third example where the individual goes to a sleeping place and overcomes an obstacle (a fallen-down Thom-Tree), we see time gain in two ways. First, it does not use some more random behavioural search for a solution as to how to overcome the obstacle, second, the particular path it chooses is quicker than the habitual way. It has a hetter chance of getting to a sleeping place before it gets tired, and there is less risk that it will fall asleep somewhere where it will be eaten by an enemy or freeze to death.

In exampel (1.4), finally, an individual plans for dealing with an approaching enemy. This planning implies a reduction in <langer. The individual will not just take its usual routine and try to flee toa tree, but takes some precautions that are more suitable (as the nearest tree is so far away). (And if there will be a fight, it may gain effort, and in the end its life, as it has a stick as weapon.)

Note that in all of the examples the relevant gains - in time, effort and security - are 'immediate' gains relating to the satisfaction of a current interest. If the immediate planner is hungry, it wants to get food quickly, with little effort and in a non-dangerous way, but is not interested in gaining effort or time 'to do something else'.

In terms of the benefits of search reduction, goal conflict resolution and error recovery, the following applies. *Search reduction* - reduction of behavioural search in trying out altematives - is illustrated in each example. In particular, in example (1.2) the system comes up with a solution which leads to a time gain. Regarding *goal conflict resolution,* consider example (1.4). The 1-Creature plans to hide from and to attack its enemy. It has two subgoals: to be hidden and to have a weapon (a thomy cane). In the plan, however, actions are ordered so that it will first get the weapon, then hide, and not the other way round. And *error recovery,* finally, is illustrated in example (1.3). When discovering that there is no clear path where it had thought there would be, but that there is an obstacle on the path, the 1-Creature stops andreplans.

4. The Value of Anticipatory Planning

Since the A-Creatures are descendants of the I-Creatures and share the environment as well as general interests, the value of their planning is best discussed by comparing the A-Creatures and the I-Creatures in the context of their respective problem solving capacities and interactions with this environment. We start by considering the examples of the A-Creatures' planning on page 122-124.

The first example, (A.1), where an A-Creature, going on an exploration tour, constructs a plan related to future needs for food and sleep, illustrates gain of time and security as well as resolution of a goal conflict. An I-Creature in a similar situation might go exploring and wander far away from food anda suitable place for sleeping. And so there is an increased risk that it will be eaten by an enemy or die from the cold in attempting to satisfy these interests. Examples (A.2) and (A.4) illustrate how an A-Creature plans and prepares to deal with potential enemies in ways that I-Creatures cannot. Using the invented method for keeping enemies away may imply a considerable gain in security in situations where it is trying toget hold of Good-Berries. Example (A.3) illustrates the resolution of a goal conflict. The A-Creature in this situation has the goals of getting shelter for cold and getting food where both goals cannot be pursued *simultaneously.* But the planner *orders* these two goals and corresponding actions *in time* - in a reasonable way. An I-Creature could not handle this situation in this way. Although it can deal with conflicts between *subgoals* relating to *one* overall goal, it cannot deal with a conflict between overall goals or goals relating to independent interests. Example (A.1) and (A.5) illustrate the phenomenon of 'being opportunistic' in planning. ('I'll take some berries from the Red-Berry-tree when passing by' and 'As I pass the Bast-Bush anyway, I can check whether it has any sprouts'.) Also example (A.6) illustrates opportunistic planning in another form. The creature plans to fix a sleeping-place when there are many sprouts on the Bast-Bushes. Doing this implies a general gain in time or effort, not particularly related to the interest of building the shelter (it will not reach this goal quicker), but if it can spend less on building a shelter, it will 'have some extra time and effort' for dealing with something else.

In sum, A-Creatures are more sucessful, thanks to their anticipatory planning, in satisfying their interests than I-Creatures. They are hetter equipped to deal with the main threats of cold and enemies. It happens less frequently that they freeze to death or that they are killed by an enemy. It is not that they have a hetter capacity to deal with *urgent* problems related to one interest and motivation, but that they can deal with immediate and potential problems simultaneously, prevent problems and facilitate and prepare the dealing with certain potential problems. Their anticipatory planning capacity implies, in various ways, *gains* in security, effort and time.

It is essential to note that the questions of *gain* (and so of adaptiveness) may be much more complex in an anticipatory planner than in an immediate planner. An anticipatory planner may not just gain time in the sense of getting food quicker and not starving, but it may gain time in a longer time perspective. It can 'have time over', plan its time so that it can obtain as many independent goals as possible, 'be efficient'. (Cf. example (A.6).) (An immediate planner can never have the *intention* of gaining time in this way.) Anticipatory planning can thus involve *decisions* of a new kind, namely, decisions about *what time perspective* to consider, how to evaluate the future in relation to the present, etc.

Another reason why the question of adaptiveness takes on other dimensions when one shall evaluate an anticipatory planner is the following. Coarsely speaking, an *immediate* planner does not plan *what* to do - what interests to deal with and what overall goal to pursue - but it plans *how* to reach a given overall goal. Furthermore, at least in immediate planners in nature, the interests and general goals are closely tied to biological hardware. The anticipatory planner, on the other hand, may be confronted with the problem of having to *decide what interests and overall goals* to pursue. Consequently, this dimension, namely, the interests and overall goals chosen, must also be considered when analysing the value of a particular capacity for anticipatory planning. Well organized planning of action - in the sense that the execution of the plan will indeed lead to the goal - but with a goal that is misplaced or inappropriate is of course non-adaptive. Thus, in the anticipatory planner there are new kinds of problems that may emerge and that should somehow be solved. Crudely put, the anticipatory planner may need an intellectual capacity not only to plan its actions to obtain what it wants, but also to *know what* it wants.

In sum, the issue of the gains secured by planning is more complex in anticipatory planning because of the time dimension and a possible new problem caused by the increased freedom to chose one's overall goals.[7]

5. Non-Adaptive Planning of Action

In many discussions on the planning of action (cf. pp.37-38) it is assumed that, in general, planning one's actions is 'a good thing' and consequently desirable. But this assumption, I maintain, is inadequate in a number of ways. It is not just that planning may be inappropriate in a situation because of *deficiencies in the planning capacity* as, for example, that the planner lacks relevant knowledge or is not capable of handling a sufficient number of means-end relations, has a memory which is not large enough, is too slow, etc, where these are *cognitive shortcomings that might be remedied.* There are also two other ways in which planning can be inappropriate.

First, planning may be inappropriate in certain situations because it is just not the right thing to do at all. It is not feasible - at least not realistically - to obtain adaptive planning by improving the planning capacity of the particular agent (for instance, by giving it more memory, improving its knowledge representations, and so on). In other words, for handling these situations adequately, one should not proceed by improving the means for generating and testing potential actions beforehand. Instead one ought rather try to deve-

---

[7]For a proper treatment of the question of the selective advantage for a particular capacity in a kind of creature, one must investigate the context in which the capacity is considered to have evolved. What were the living conditions of these creatures? Did their planning capacity make it possible for them to salve problems better or engage in new important activities, and so on? Were there any changes in the environment; in the challenges and opportunities it gave, that may have implied a 'pressure' towards the evolution of the capacity? A real investigation of this problem concerning the evolution of more or less sophisticated planners demands a thesis of its own. In the chapter on human planning, I will give a few examples of tasks where a capability for anticipatory planning could have rendered a selective advantage for the hominids; namely, for the use of fire and for the hunting of larger animals.

lop the means for producing adequate actions moment-to-moment, for improvising, using routines, and so on. (Cf. planning is not the solution to all action problems.)

Second, planning may in certain situations be inappropriate *even though* the planning competence is sufficient and adaptive planning objectively feasible in the sense that thanks to planning, time, energy and so on may be gained and goals easier reached. There are, namely, certain kinds of goals and interests that restrict the desirability of planning. In particular, this applies in a social context.

In sum, we have the following three kinds of situations:

- The planning capacity is not powerful enough, but it is possible to improve it.
- Planning is not a solution to look for at all; it is not objectively feasible.
- Planning is objectively feasible but in conflict with other indepedent values.

In traditional approaches to planning it is only the first kind of non-adaptiveness, but not the latter two kinds, that is regularly recognized. This is not so surprising. The classical AI conception of planning and plans, recall, (see p.38) is that *plans* underlie all behaviour, in the sense that all observable structure in behaviour comes from a plan with the same structure. Plans are the universal formulas for behaviour production. Thus, if something goes wrong in behaviour - i.e., if some behaviour is maladaptive - there must be something wrong with the plan and the planning. Perhaps the planning capacity is not *powerful* enough, and so we should try to improve it, for instance, by giving the planner more memory, a hetter representation of its world (more knowledge), hetter ways to gain access to its knowledge, a capability to handle longer sequences of actions or a greater number of goal-means-relationships, etc. According to such a view one cannot, of course, question the desirability of *planning* as such. But this, I maintain, is not how planning should be conceived. Rather, *planning* is *one* of a set of *alternative* activities that can be undertaken in a particular situation. And in some situations it is preferable to leave out planning and rely upon other behaviour-regulating strategies; either because planning will not lead to any objective gain for some goal attainment, or because there are independent criteria and values that planning cannot meet. These are the two 'alternative kinds' of inapropriateness, and I will now discuss them in some more detail, starting with the first one.

In parallel to the discussion about the value of planning on page 157-158, the general conditions under which it may be hetter to refrain from planning are when planning will not imply any gains - in time, in effort, in security or some resource that is scarce or valuable - that hetter enable the planner to reach one or several of its goals than if it uses a routine or habit, just tries things out or 'goes with the wind', etc.[8] Furthermore planning is not desirable if the *costs* of planning are greater than the possible gains. However, this is something that may be hard to estimate.

In general, the planning of action is not 'objectively feasible' if

---

80r where the likelihood that it will do so is small.

• The planner cannot predict the situation in sufficient detail (either because its knowledge is *not updated,* or because it *cannot gain* the required knowledge).

• The planner cannot control relevant aspects of the situation sufficiently.

• The planner does not have enough time for planning, cannot plan quickly enough.

If a system engages in planning under such conditions as these, I will say that it engages in *non-adaptive planning.*

On page 114 I presented one example of non-adaptive immediate planning; a person without updated knowledge about the restaurants and communications in her old hometown who sets off to plan where to go to have lunch. Or consider the situation that you are about to have lunch somewhere and the weather seems very unstable. It is then not appropriate to make a plan to have a picnic if an essential condition for this is that it does not rain. The weather, certainly, is something of which one is not in control. For another example, it is non-adaptive to sit down and think through what to do or how to do something in situations where there is an urgent problem demanding quick action and where planning will rather imply a loss of precious time; for instance, when an 1-Creature is facing an enemy. The 1-Creature in example (1.4), page 94, on the other hand, *has some* time to think about what to do. (Yet also in this case it should think quickly.)

When turning to *anticipatory planning,* the issue of adaptiveness vs. non-adaptiveness is, as I have pointed out, more complex. Just as there are more ways in which one may gain (and intend to gain) something by anticipatory planning, there are also more ways in which it can go wrong and be non-adaptive. I will give some indications and examples of this.

Due to the longer time perspective that can be involved in anticipatory planning, in contrast to immediate planning, and to the greater number of *possible* independent actions, interests and problems considered, there is an increased risk that one plans actions that will never be realized and actions that may not be possible to realize (like if one plans a skiing trip that can then not take place because one becomes ill or because of inadequate weather conditions; or if one plans, in the morning, to spend the evening in a particular way but when the evening comes there is a new unexpected opportunity; or there is an obstacle to one's plans, no means for transportation, for instance; or one is not 'in the mood' for doing what one had planned). In the general case, adequate anticipatory planning, as compared to immediate planning, requires more knowledge. There are more facts to take into account, and all these knowledge items are prone to the risks of insufficient predictability and controllability.

By anticipatory planning one may *gain* time, resources and effort in a way that can not be obtained by immediate planning, namely, in the sense of saving resources for the satisfaction of other independent interests and for future potential interests. (Cf. pp.161-162). But anticipatory planning can also entail a *loss* of time and effort. In line with the foregoing paragraph, the anticipatory planner may spend time and cognitive effort on potential problems and alternative plans, etc., that will never become actual or realized. Furthermore, anticipatory planning may involve preparation and arrangements for predicted problems, where the problem never becomes actual, or, when it does become actual, there are much easier ways for dealing with it. (For instance, a person is at a conference and realizes that she has forgotten her money at the hotel and is concerned about how to get some lunch. She thinks about whether she can go back to the hotel and how or whether to ask someone to lend her some money. But then before lunch, she is informed that lunch will be served at the expense of the conference. In a worse case, the person is so occupied by thinking about how to solve her problem that she misses the information.) And so the time and effort that the anticipatory planner has spent on thinking about and preparing for this is a waste. It could have been used for other things.

What is crucial here is that planning is *not for free*. Planning is undertaken *at the expense of* acting and perceiving (cf. p.99). First of all, the individual, *while planning*, does not fully engage, perceptually and behaviourally, in the present situation. Even though the planning may not be undertaken at the expense of engaging behaviourally in any *immediate problem*, it may occur at the expense of, say, playing around, exploring and discovering. However, it might be *more valuable* to engage in *these* activities - which are also anticipatory activities - than to engage in anticipatory planning. The point can be put as follows. If a creature has a tendency to engage in anticipatory planning and fantasizing to a great extent instead of playing exploring, etc., this may be detrimental in the long run, even though it is not obviously and immediately detrimental. An anticipatory planner is necessarily a creature that does not have to be constantly involved in a serious struggle for life; there *is time* for considering the future, thinking, maybe 'daydreaming' and indulging in fantasies. Yet it is important that this activity is balanced with and fitted properly in with other activities. And it is important that it is engaged in roughly 'at the right moments of time'.

Furthermore, anticipatory planning may be non-adaptive in that the *overall goals* that the planner selects are inappropriate. With the loosening of motivational forces, recall, anticipatory planning allows the construction of interests, goals and problems that are not immediately related to biological interests.[9] (Cf. p.162.) The system may be *planful* in relation to an interest that is not biologically relevant. I shall give examples of this in the next chapter on human planning.

Let us now briefly turn to the second 'alternative way' in which planning can be non-adaptive. This involves the introduction of certain kinds of interests and goals that can restrict the usefulness and desirability of planning, even if planning is 'objectively adaptive' in the sense that it implies a gain in effort, time or resources relative to certain interests during a certain period of time. In particular, the introduction of a *social context* with social goals and values can produce a certain complexity. I will discuss this more extensively in the next chapter on human planning. Here are just some examples (concerning human planning): A person makes a plan as how to let a guest stay with him but to avoid his guest as much as possible because he does not want to see her now but thinks he may *need and want* her around next week. (OK, I will put up with this now because next week I may want some company, so I will keep this relation going.) This kind of planning may, however, strike one as non-desirable and inappropriate; there are restrictions on making plans that involve other people, to *plan* human relations. As another example, a plan made in advance to seduce someone may work out, yet this may be inappropriate becasue, again, this concerns the intricacies of human relations, and we also

---

9 This is not something that *has* to happen, but it can do so.

puta value on acting spontaneously. (Note that these considerations bring in the social and the individual experiential perspectives on value.)

In sum, an agent sometimes does hetter to leave out planning. Toere are situations when one ought not to try to think beforehand and predetermine a course of action in order to reach some goal, but where it is preferable to just rely on habits, or to flow with the tide and take opportunities, or to improvise and try out actions and so on, to reach one's goal(s). There are basically three general reasons for this. First, the agent's planning capacity may be insufficient; second, adaptive planning - that is, planning that implies gains of some kind for the reaching of one's goals - may not be realistically feasible; third, even if adaptive planning is (realistically) feasible, thre can be other kinds of losses eaused by the planning.

Note that I have mostly spoken of situations where it is in-appropriate or non-adaptive to *plan at all.* But of course one may also find *non-adaptive planning* in the sense that it is the wrong *kind* of planning in a particular situation. For example, when you engage in *detailed* advance planning in a situation where the number of uncontrollable events is large and you have too little control, but where it could yet be adaptive to engage in short-term planning with multiple options left open; or when the planner uses too much top-down thinking in a planning situation where it would be more appropriate for it to be opportunistic in its planning; or if the planner goes through too many altematives (or to few), or uses inappropriate sources of information in planning, and so on.

Another remark to be made is that adaptiveness never only concerns how and when the planner plans but also how plans are executed, how flexible the planner is in following a plan; for instance, how much attention he pays to opportunities not considered in the plan (see title page). (This, again, may be a more critical issue for anticipatory than for immediate planning.)

To end this chapter, let me add some brief comments concerning why non-adaptive planning may appear in evolved biological creatures. First, it is not necessary that a capacity is adaptive in *all instances,* or in all of its applications, in order to be selected. Thus, in a system with a planning capacity that is *generally adaptive* - and therefore selected - there is room for local non-adaptiveness. Second, there is also the possibility of more extensive and serious non-adaptiveness if a system's environment differs too much from the environment for which the planning capacity was selected and is adapted to. The first kind of non-adaptiveness may occur within relatively constant environmental conditions, whereas the other is due to environmental change. I will return specifically to the second type, when I discuss human planning. As a general point, any capability may, once it exists, be used for other things than those for which it was selected - and this can be adaptive or not. Planning may go wrong or be maladaptive when a planning capacity is used in the wrong situations or on the wrong material.

12. ON HUMAN PLANNING

1. Introduction

The underlying motivation for writing this thesis has been my interest in the role of planning in the activity and life of human beings. Throughout the thesis I have now and then referred to planning in human beings (in particular, in the introductory chapter in part I, in the chapter on anticipatory planning and in the preceding chapter on the evolutionary value of planning). In this last chapter, however, I will speak more explicitly of the role of planning in human activity and of hu_mans as planning creatures. I want to emphasize that much of the discussion in this chapter is speculative and the proposals I make tentative.

In brief, my conclusions will be that *human beings* are *anticipatory planners* - and the only anticipatory planners in nature; that there is much human planning that is *non-adaptive;* and that man is a *'creature in time'* with its positive and negative implications (see pp.138-141). I argue forthese theses in the next four sections, and finally there is a section on the ontogenesis of anticipatory planning in humans.

2. Human Beings as Anticipatory and Social Planners

Human beings are anticipatory planners. Let me start by saying what I do *not* mean by this. I am not saying that all human behaviour is preceded by anticipatory planning (or by planning in general). Planning and plans, in my conception, are not sufficient to produce behaviour. And it is not the case that all *structure and adaptiveness* we see in human behaviour is due to a corresponding structure in an underlying plan. Regarding the *importance* and the *prevalence* of planning, and of anticipatory planning in particular, this differs, I believe, considerably between individuals, and probably also between cultures. Yet I *do* maintain the following:

• Humans, but no other creatures, are capable of anticipatory planning - that is, the predetermination of actions explicitly aimed at achieving some goal that relates to future problems, interests and needs - and anticipatory planning is an essential part of life for many human beings.

• In circumstances that are not too extreme or deviant, a human being will develop the necessary preconditions for anticipatory planning: A suitable motivational structure and some form of time representation. She will develop a conception of a past, a present, and a future, and function motivationally so that she will to some extent be conemed about the future and her potential interests. But exactly *how* this capability is developed and used in an individual also depends on the individual's cultural and social environment. Toere are circumstances w ere it has hardly any chance to develop at all, like in environments where humans are struggling for their immediate survival.

Let me begin to comment on the emergence of anticipatory planning in nature. First, human beings are obviously capable of anticipatory planning, and there is no evidence that any other creatures possess this capacity (cf. section 5.7). This leaves us with the hypothesis that this capacity emerged at some point in the transition from apes to humans. And there are some indications that this in fact was the case. One sign of the emergance of anticipatory planning is the *use offire* for which there is evidence in early man but not before this. To keep a fire going requires collecting fuel *before* one gets too cold. (See illustration below.) Moreover, the *hunting* of */arger* animals assumed a great importance in early man or in the hominids. Hunting, tracking, killing and preparing the meat of larger animals inevitably involved the cooperation of a large number of individuals. And so presumably, groups of humans or prehumans worked together, planning and cooperating in order to snare and share the meat of the bunt. This kind of bunting is not something one sets out to do when driven by immediate hunger. *Graves* of various kinds are interesting as well, as they indicate a representation of one's own future - a sophisticated time representation. Notice here a difference in the way chimpanzees behave. Although chimpanzees react violently when another chimp of the group dies, they rather behave as if they would expect the dead member to rise and come along. Toere are no signs of an awareness either of what happens to the observed individual nor of a transfer of the event to their own case. (Cf. Bischof-Kohler, 1985.)

Illustration 11. Use of fire

These are some indications of a capability for anticipatory planning.The more fundamental question, of *how* the preconditions for a development of anticipatory planning were developed (a suitable motivational structure anda time representation) is perhaps the most interesting - but also most intricate - question. I will not attempt to speculate about this, except by saying that hominids or early men presumably were increasingly 'liberated' from a constant involvement in a struggle for survival. They had a relatively long life and they regularly engaged in exploratory activities. These are suitable conditions for transcending a mere satisfaction of current needs and obtaining a recognition of oneself as a separate *entity* that *has* various needs, and a future.

I have focused most of this thesis on the case of *one* individual planning *its own* actions in contexts or environments that are *non-social.* To discuss human planning, however, the social aspects of planning must be taken into account. First, in planning their own actions, human beings often take into consideration the motivations, intentions, beliefs and actions of other agents. In other words, they consider other agents as agents. Second, people engage in collaborative planning. (For instance, some people who travel together can construct a plan together as bow to find somewhere to sleep the coming night, or two bankrobbers can together produce a plan for robbing a bank.) Third, people make plans for other people' s actions. (For instance, a figureskating instructor plans the training program as well as the particular dances for her trainees.) By *social planning,* I mean the planning of actions that involves other agents in one or several of these ways.

There are, I believe, certain interesting respects in which human anticipatory planning is influenced by her social context. The presence of social structures and forces (families, schools, hospitals, health controls, pensions...) *increases* as well as *diminishes* the need for *anticipatory planning in individuals.*

First, plans and actions concerning future interests of yours can be *made for you.* (Parents may set up an account for their child to attend college or to buy a flat when the time comes that it will need one. The leader of a camp may tell the campers what they need to take along when going on a hike. Mummy or daddy may see that there will be some food on the table tomorrow even though the shops will be closed.) And, as soon as there are some kind of social groups, there can also be *anticipatory behaviour and activities* where anticipating future needs and interests need not be tightly linked to any *explicit reasoning of individuals.* One simply *does* certain things in a certain way. For instance, one regularly gathers wood for making fires, one collects and stores food, one regularly goes to the shop, one builds houses or shelters in a particular way so that they will withstand various kinds of weather, even if it is warm when they are built, etc. These and other factors diminish the need for anticipatory planning in individuals.

On the other hand, there are created *new* opportunities for *individuals* to engage in anticipatory planning within the framworks of social structures of various kinds. First, there is in general an increased freedom from dealing with momentary survival that follows from the existence of social structures. Second, in societies we find phenomena such a ways for storing food (caves, consetvation, refrigerating, etc.), systems for buying tickets *in advance* and booking things, sales of winter clothes when winter has passed, cultural knowledge and norms of bow the life of a person ought to progress (obtaining certain positions, getting married, having children, makinga career, and so on) and such phenomena create more opportunities for anticipatory planning.

A third way in which anticipatory planning is affected by social structures concerns the cultural or social *transfer* of the activity. Anticipatory planning may be taught or more implicitly transferred. You see or listen to others making plans and may learn, for instance, to plan a voyage; to think about what you will need at various points during the trip, or you may learn that it can be advantageous or even necessary to think about buying tickets in advance, and so on.

It is important to emphasize that from the fact that there are social structures, it does *not* follow that all structure in human behaviour can be explained by such societal and cultural forces and structures, to the extent that there is no need to assume that some behaviour is structured by *individuals* that *make plans* for their own actions (and sometimes forthose of.others). I maintain, to the contrary, that explicit planning in individuals playsa role in human life and activity. Individuals do *generate* some *mental representations* of potential courses of actions, which structure their (future) behaviour. Human beings *have* some conception of a future and of *theirfuture needs and desires.* They consider this future important and *think of it* to some extent, thinking about what to do during the coming weekend, thinking about, looking forward to and maybe preparing tomorrow's dinner, thinking about whether to phone one's mother in the evening or rather tomorrow, thinking about what clothes one will need on a certain trip, and so on. By stressing this, I oppose what I have called the 'new wave approach' to planning (cf. p.39), which can be found in AI and in anthropology and to some extent in psychology, where one seems to say that there *is* no individual planning in this sense, or at least that this kind of planning plays only a marginal role in the structuring of people' s actions. According to this trend the *plans* that exist and are used are *collective and somehow externally represented plans;* maps, recipies, and so on. Toere are the views of L. Suchman (1987) (whom I mentioned on page 39) that seem to have a strong influence. Suchman stresses that *all human action* is basically *improvised* and not planned. It is improvised, though, within a social framework, and this is where the *collective plans* enter. These plans, however, do not predetermine any action patterns by specifying bow one shall proceed. Instead they are available to individuals as criteria for judging, *during improvisation,* what one has (just) done. They are criteria of success that one may bump against during action and are used for judging progress step-by-step. I also mentioned, as representatives of 'the new wave', P. Agre and D. Chapman (see p 40). They argue that it may be a mistake to place cognition *within* individual minds. Cognition, they propose, is something that takes place in the interaction between an individual mind and the world, and the human world is a social world. *Planning,* according to this view, is a social phenomenon. Planning and plans are communication in natural language. Plans are *external,* not internal or mental, symbolic structures. (See Agre and Chapman, 1987.) This fits in with the trends in antrophology that stress a view of cognition as primarily a collective product, not tied to any individual minds (see Geertz, 1973).

The good thing about 'the new wave', to my mind, is that it isa counterweight to an earlier dominating view, primarily in AI, that ascribes plans, in the sense of internal representations in individuals that predetermine or structure behaviour, an enormous role. According to this view, the construction and use of plans is the general formula for behaviour production. Plans are what causally engender behaviour, or, in other words, planning and plans are sufficient to drive a robot. (As I have pointed out, this view uses a wide notion of plan, where the origin of a plan is irrelevant. This notion seems to me to have its roots in Miller, Galanter and Pribram's classic *Plans and the Structure of Behaviour, 1960).* A danger with 'the new wave' arises, however, in the tendency to go too far towards *the other extreme* and to deny any importance of plans in the sense of internal representations in individuals that predetermine action, or even to deny the existence of explicit planning in individuals. But a primary aim in this thesis has been precisely to find the place for this phenomenon.

Note that I do not question that the ideas and conceptions of 'the new wave' are valuable for dealing with certain *aspects* of planning (cf. p.40). I do not believe, however, that they are adequate for a *general* treatment of the character of planning in humans and other creatures. I thus oppose a tendency to view psychological phenomena in humans as fundamentally grounded in 'advanced' *cultural and social structures,* requiring linguistic capabilities amongst other things. I beleive there are principles to learn about by studying more primitive or basic forms and contexts of psychological phenomena.

3. Cultural and Individual Variation in Human Anticipatory Planning

I think it is fair to say that planning in general, and anticipatory planning in particular, is an essential feature of *modern Western culture,* and that Western society toa large extent is characterized by this.Yet the prevalence of planning, and of anticipatory planning in particular, may vary between cultures. For instance, cultures seem to *value* planning differently. In modem Western society, planning seems to be seen in an unequivocally *positive* light, whereas it seems much less so in some other parts of the world. (J.

Goodnow, 1987, gives an example of a practice on the Philippines, namely the practice of 'barging ahead' or 'acting anyhow', that is, acting without a plan that specifies some way toget to ones goal. This practice is *highly regarded* and thought of 'as a good thing'; it would not be so, she contends, in the United States. *(Ibid.,* p.194.))

If my ideas about the role of an explicit time representation in anticipatory planning are correct, the role of *anticipatory planning* and how it is valued is probably intimately tied to the particular character of the time conception that is promoted by a culture and society - that is, how a mimimal time representation is elaborated in a certain culture (for instance, bow much emphasis there is on time as a scarce resource, on the saving of time, whether there is emphasis on linearity or cyclicity, and so on). This will influence bow individuals develop and elaborate their conception of time.

In this context I want to mention Whorf' s famous study of the Hopi Indians. *ff* it is true that their time conception differs from that of the Western society in the ways Whorf relates, we should expect differences concerning anticipatory planning as well. Assume, for instance, that it is true that the Hopis conceive of ten consecutive days rather as going to see *ane and the same person ten times* consecutively than as seeing *ten different people* consecutively, which may be nearer a Western conception (Whorf, 1956, pp.139-140.) In this assumed Hopi conception, there is an emphasis on *continuity* in a way that could be expected to have consequences for planning and anticipating. According to the proposed conception it is in a sense always 'the same day' we are dealing with, and so we can prepare for tomorrow by acting today, and we can know things about tomorrow by finding out about today. It could be expected that this implies a stress on preparatory activities but that there will be less stress on possibilities of starting afresh tomorrow, beginning anew, meeting new problems, and so on. *Both* the presumed Hopi and the presumed Western conceptions of days make it possible to *compare* today and tomorrow and regard today as well as tomorrow as *important,* but the details of the comparisons are different. The Western conception leads to more stress on the *formal* likeness of yesterday, today and tomorrow, as *equivalent units of time,* formally commensurable.

Note that I speak of the possibility that the time conception of Hopis is *different* from that of Western society. I am *not* considering the possibility which has at times been suggested (first by Whorf himself) that the Hopi indians should be described as having *no* conception of time at all - as their conceptions are so different from Western conceptions that it is misleading to use the term 'conception of time' in both cases. My hypothesis is that both Hopi Indians and members of modem Western society *have, as human beings,* an explicit time conception with some notions of a past, a present, and a future and some ways for comparing them. (For arguments supporting this view see Bach, 1981.) How this 'minimal' time conception is then further structured and elaborated may well differ- and seems to do so - between Hopi and Western cultures.

So far I have spoken of *cultural variance* regarding anticipatory planning. But also *within* cultures one will find differences between individuals as to the importance and prevalence of anticipatory planning. Even within modem Western society which is a society that clearly puts a premium on advance planning, we find a wide spectrum. At one end, one finds the pronounced anticipatory planner who is continously taking potential needs and problems into account (who will always bring extra clothes along, who even when going to the beach in the hot sunshine may think about the possibility of staying and that it may get cold, who when buying a flat will primarily think about what he will need in a few years time and adapt his criteria and actions to this, and so on). At the other end, one finds the 'non-anticipating' individual (who will almost always have an empty fridge, who will not buy, say, an extra beer for tomorrow but only go and buy things when he wants to have them, who may go on a safari in the desert with no water, who may have considerable difficulties handling a credit card, and so on). [10]

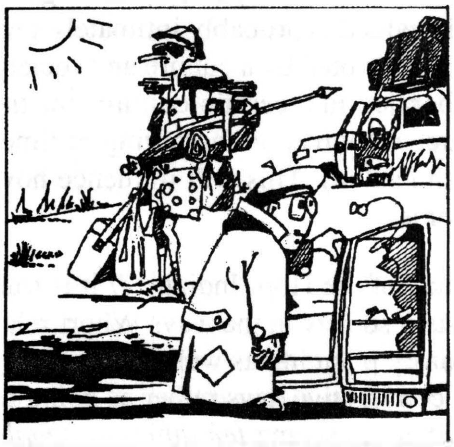


Illustration 12. Anticipating and not

Summing up the two last sections, humans are anticipatory planners, and rather sophisticated anticipatory planners. The capacity for anticipatory planning did, I believe, render humans advantages over other creatures in the course of evolution. And it is easy to find examples where individuals who engage in anticipatory planning in a situation may do hetter than individuals who do not. (Cf. the examples above.) But there is also much anticipatory planning in humans that is *non-adaptive,* which will be the topic of next section.

4. Non-Adaptive Anticipatory Planning in Humans

In the chapter on the evolutionary value of planning, I discussed non-adaptive planning in general, and in this section I will tie onto this. [11] One reason for writing so much about non-adaptive planning is to counterbalance the tendency in many studies of planning which unquestionably regard planning as something unequivocally desirable. But there is much human planning, and human anticipatory planning in particular, that is non-adaptive and not desirable. This holds within the context of biological or evolutionary value, within the context of social value as well as within the context of individual experiential and reflective value (cf. p 159). In general I will discuss the non-adaptive aspects of planning from the evolutionary or biological perspective unless otherwise noted, but I also bring in the social and experiential perspectives.

Fir t, there is the risk of overengaging in planning (cf. p.139). This is exemplified in human planning which, it seems, costs more than it is worth, as the planning, and antici-

10 A question that one may pose is to what extent education and familial influences determine the development of individual differences.

11 Note however that this discussion is tentative and far from an exhaustive discussion of non-adaptive human planning

patory planning in particular, occurs at the expense of other things, like dealing with the current situation. (Cf., for instance, Anita in the first example, on page 3-4.)

In some instances, it seems, planning in humans (in particular anticipatory planning) has become an end unto itself. This may have to do with the fact that anticipatory planning can be partly self-sustaining (cf. pp.139-140). When planning is removed from action and there is no present time perception or interaction with the environment that directly relates to the content of the plan, the *motivation for planning* must lie elsewhere. It may lie simply in the activity of planning itself. Anticipatory planning must to some extent be reinforcing and enjoyable as such. An individual may then simply enjoy making plans for the future, even when the likelihood that she will act according to them is minmal or even when she *knows* that these are actions that she will never carry out and so on. (Cf. pp.139-140.) This liability may lead to non-adaptive planning.

Furthermore, in anticipatory planning, in contrast to immediate planning, the plan execution is often temporally non-contiguous with plan formation. This implies a difference conceming the possibilities of *learning how to plan.* In immediate planning it never takes long between planning and plan execution, and thus there is some direct feedback as to whether or not it was a good strategy to plan in this kind of situation, whether or not it was a good plan, etc. In anticipatory planning, though, there is often no such direct feedback and sometimes no feedback at all. Human beings can, I believe, easily develop *planning habits* that are inappropriate, but yet keep them, like, for instance, if one makes long term plans for the coming holidays or weekends and always tries to fit in too much, or if one continues to plan one's days of work in a particular way even though one knows that those plans are not very helpful.

We also find the danger of anticipatory planning together with insufficient flexibility in plan execution. This situation occurs, for instance, when someone makesa plan in relation to some goal which later turns out to be not so important and which perhaps even excludes something that is more important, and yet he acts according to his plan, 'because he had planned this'. The planner may have misjudged his future needs because he did not know or think sufficiently about them.

On page 165, I also mentioned that the planning of action can be non-adaptive because the *overall goal* of the plan is of questionable value. This is definitely often exemplified in human anticipatory planning. Anticipatory planners, recall, can be subject to this problem because the more firmly biologically grounded forces for directing action are dampened- which is a step that *has* to be taken in order to make room for anticipatory planning. Now, as long as the planner has a restricted number.of problem domains - a limited number of needs and goals - to deal with, it is not so problematic. Human beings, though, seem to have a multitude of independent interests and motivations, many of them culturally and socially created. In particular, the fact that human beings have an explicit time conception, can shape new interests, such as interests in property, interests in revenge and new interests in reciprocity. What this means is that some human anticipatory planning must also crucially involve figuring out and deciding *what goals* to pursue. Consequently the situation can arise where planning is successful - in the sense that the execution of the plan will indeed lead to obtaining the goal in an advantageous way - but where the goal is in some sense misplaced or inappropriate or just unimportant. For instance, people can be *planjul* with *resources that are not scarce* - counting the seconds while on holiday, or the truly rich keeping track of small amounts of money. The planning may be successful, and yet one may ask about the value of the goals reached (from a biological and social as well as from an individual experiential perspective). What seems to happen is that the value of planning, in that it minimizes the waste of resources, is 'cut loose' from an appropariate context. One forgets to think about what one really gains; what the *overall goal* is that shall motivate the *minimizing of waste* or the *efficiency.*

I also promised in the preceding chapter to say more about non-adaptive planning that is due to changes in the environment of some creatures. I will do so by discussing the human planning capacity in modem Western society. The basis for my discussion consists of the following considerations. The human being is an animal that has evolved over millions of years and is an adapted member of an ecological system. During 99% of human history, the circumstances under which humans lived were probably relatively invariant. Their life form was characterized by the following features. Humans lived in *small groups.* They had no permanent dwellings but undertook seasonal wanderings within a limited area. They gathered plants of various kinds and hunted small and large animals. There was cooperation and food sharing. There was hardly any storage, and personal belongings must have been limited to that which an individual could carry along. Thus, there was no food-production, no domiciles and no conglomerations of larger groups of human beings. (See Bischof-Kohler, 1985.) The comparatively *recent* agricultural, technical and cultural evolution has implied substantial changes in the form of life and living conditions for human beings. But there has not been any corresponding change of human fundamental or basic biological-ethological functions: human behaviour control systems - human ways of perceiving, feeling and reasoning - are adapted to the just described conditions of life. I will make some remarks on what might therefore be the results when the human planning capability and ways to make decisions are applied in present Western society. I base this discussion upon G. Goude's paper "Man - a biological being or a technological mistake?" (1977).

Planning crucially involves making *predictions.* Human beings have certain inclinations that influence and limit the ways in which they make predictions. For instance, humans typically utilize representativeness instead of probability to estimate how probable different outcomes or events are. This, as Goude points out, is an adaptive strategy, when man acts in his own adjacent and familiar environment, but not necessarily if he tries to deal with larger and less familiar contexts. Goude further refers to G. Ekman's studies (1965, 1970, 1971) that show that personal involvement in an event is a *rapidly* decreasing function of the distance in time or space to that event. Such a way of functioning seems suitable for the life form described. But it may not be adequate that our evaluations of the importance of problems and goals is coloured in this way when the conditions under which we live are such that we get information about events *distant* in space and time, and that we sometimes make decisions and plans where such events shall be considered. *(Ibid.,* pp.5-6.)

These are some examples of functions - involved in planning - that may be adequate, or at least not maladaptive, for the life form described, where humans lived in 'well-known' surroundings and in small groups. Yet, planning concerning technical, social and cultural issues in large contexts and for long time spans may turn out less well because of such functions. The kinds of mistakes due to limitations as the ones mentioned may be relatively *harmless* as long as the system is dealing with a narrow context in spatial and temporal measures (and in complexity), but can be amplified when the context is broadened. One example may be the human tendency to overestimate bow much can be accomplished duringa certain period of time, documented in everyday experience as well as in experimental studies such as Hayes-Roth and Hayes-Roth's experiment (1979) where people are asked to make a plan for doing errands during an aftemoon. Another may be our limited ability to keep different items in mind simultaneously (see Miller, 1956; Ellis and Hunt, 1983). This may not be much of a problem as long as a planner's environment is not too complex and the number of possible independent goals limited. But in a more complex environment, as that of modem human beings, where a planner may have to handle many independent goals and items of information related to them, this may be a severe limitation. It is certainly no wonder that *planning in /arger contexts and on a societal leve/* can be difficult. Goude, in the mentioned article, points at three possible strategies for dealing with this. The first is to try to let external tools, computers for instance, take over some planning and decision functions. The second strategy is to try to 'improve' human capabilities by teaching and training planning and decision making. The third strategy, the one he favours himself, is probably the most original and least tested strategy; namely to *adapt* social and technological systems to *human cognitive limitations* and to *adapt* decision situations to human competence as given by biological evolution.

Finally, since much planning in humans is social planning I will say a bit more on non-adaptive human planning *in social contexts.* The main part of this material comes from J. Goodnow's paper "Social aspects of planning" (1987). Goodnow first proposes to generally describe the planning of action as consisting of the *movement ofpieces* from one state to another. One alters their places in a sequence, tries to compress the time usually allotted, combines pieces that are usually separate and uses old pieces in novel ways, etc., all with the aim of progressing efficiently towards a goal. The problem in social planning is that now some of the pieces to be moved around on the planning board are *other agents* (with a status similar to the planner's), and these are pieces that may be *particularly difficult to predict and control.* Therefore, there are, first of all, cases where the planning of action just is not feasible - because of the unpredictability of and lack of control over other agents. But secondly, also when planning *is* feasible it may be undesirable. The reason for this is that the social context may introduce new kinds of criteria for what is desirable conceming the planning of action. I touched upon this in the chapter on evolutionary value, where I also gave some examples of this (planning a seduction, making a plan to keep some company). (Cf. p.165.) The following quotations from Goodnow's paper, of some overheard utterances, are additional illustrations: "He treats going to a restaurant like a military campain"; "I know I'm late, but I'm going to a party, I'm not catching a train"; "It's weird, she has this schedule she has to check to make sure she says three nice things a day to the kid." *(Ibid.,* p.179.) Goodnow also relates a story of a woman who gladly reports that she has managed to fit in about ten social calls while visiting her mother. When she reports this and asks "Wasn't that efficient?" the reply she gets is "Much too efficient." *(Ibid.,* p.185.)

In general, I conclude with Goodnow, the existence of proper interpersonal relationships will often call for the suspension or limitation of planning. Furthermore, there is a value in spontaneity, 'being carried away', 'acting out of feeling', etc. And thirdly, from the perspective of experiential value, some actions may be *boring* if they are planned (too well) in advance.12

Summing up this section, it is clear that planning in humans is not always a good thing. Toere are plenty of examples of human planning that for different reasons can be regarded as non-desirable.

5. The Creature in Time

Human beings are - sophisticated- *anticipatory planners.* I indeed venture sofaras to say that this capacity - with the ability to represent future, potential problems and needs and to care and think about them - is a *specifically human capacity* and that it is essential for the human condition. 13 It is a capacity which, I believe, also goes along with some other characteristics that we deem characteristic of human beings. It allows them to have a larger action spectrum than other creatures since they have less forceful and determinate motivational forces and since conceptions of the future may influence individuals' actions. Human beings even have the perspective of death and a related existential anxiety, which is often pointed out at as what distinguishes humans from other animals. Human beings can think of themselves as getting older, and can think of - and even plan - their own death. (Furthermore, *planning* can possibly be used as a means for keeping death at a distance. As long as you have plans for your future there is something that separates you from death.)

On page 138-141 in the chapter on anticipatory planning, I spoke about *creatures of time,* that is, creatures that on a contemplative/reflective as well as on an experiential level are strongly influenced and marked by their time representation. Humans are, I contend, such creatures - *creatures of time.* The way they perceive a present situation, think about it and act in it, is more or less strongly coloured by their representation of their own past and future. Pure representations (thoughts, images) - of past and future potential events - can have strong effects on current interpretations and experiences. The positive and negative utterances conceming planning and thinking about the future that I related in the introduction (p.5) can, I believe, be understood against this background. Planning may be *experienced* as negative by an individual when it occurs at the expense of present perception, action and experiences and when it too rigidly restricts his acting. Anxiety conceming the future, related to anticipatory planning, can also be placed on the negative

12In this discussion of social planning I have focused on possible *non-adaptiveness* of planning in a social context. It should, however, also be pointed out that planning, and anticipatory planning in particular, undoubtedly may be adaptive or valuable from a social perspective. Anticipatory planning can make cooperation and coordination of actions possible, which may lead to a better reaching of collective goals. (Think of the planning of a bunt, a meeting, a course, a party that you prepare together, and so on.) Collective planning - planning together - moreover, can lead to group cohesion and to important social experiences, in that a sharing of plans may also involve a sharing of dreams and ideas and a sharing of responsibility.

13Note that I am not saying that humans are the only creatures that plan but that there is a characteristically - unique and universal - human planning capacity. Similarly I would say that it is not so that humans are the only creatures that have a language capacity but that there are characteristically human linguistic capacities, and that humans are not the only creatures that can laugh but that there is a specifically human laughter. (These are some of the capacities that have been suggested to be 'defining characteristics' of man or human nature.)

side. On the positive side, from an experiential point of view, there is the link to imagination and daydreaming, that many people experience as something positive. (And this is probably also positive from an evolutionary point of view, even if there is the question of *balancing* intemal and *external* exploration and play that I have mentioned, p.165.)

One way to illuminate the positive and negative sides of planning from an *experiential* view is to consider how *planned action* can be characterized *in contrast* to *non-planned* action in everyday conceptions and speech. This can be done in (at least) three ways:

(1) An action is planned and is *not* just done out of routine or habit.

(2) An action is planned and *not* performed without thinking. It is not rash and hasty, ill-considered and not reflected upon, but there is some thought behind it.

(3) An action is planned and is *not* spontaneous and improvised.

The first two characterizations point to 'the positive side' of planning. But the third characterization gives another side or view. An action that is planned is *not a* spontaneous action. It is not created on the spur of the moment, as a surprise to oneself and others. It is not an action out of feeling. This is obviously something we do not want to miss out on.

6. *The Ontogenesis of Anticipatory Planning*

When one presents hypotheses about evolution or phylogenetic development, as I am proposing that human beings but no other creatures are anticipatory planners, it is close at hand to ask about ontogenetic development. Are children anticipatory planners? Are there any parallels to look for in ontogenetic and phylogenetic development?

These questions have not been adressed directly in the literature, for the simple reason that the distinction between immediate and anticipatory planning, as far as I know, has never been explicitly considered. However, I have, in the literature that I have read on the ontogenesis of planning of action in general, found some indications of interest. I will relate some of the findings that Kreitler and Kreitler (1987) present in their article "Conceptions and processes of planning". When probing what children - aged from 5 to 11 years - themselves *understand by planning,* the following trends show up. Younger children (5-7-year olds) consider planning to be something one does primarily in regard to regular daily activities such as eating, going to bed and dressing in the morning. It should be noted that all these relate to relatively immediate needs. Furthermore, planning is considered as "applicable only if there are immediate benefits for the planner" *(ibid.,* p.216.). From about age 7, however, children will also speak of planning as something that is used to attain *control over oneself* (and of others); one may plan one's homework, plan some errands, make a plan for actions in order to maintain one's health, etc. Asked about the *purpose* of planning, 5-7-year olds answer that it is necessary to plan certain actions to be able to perform them. They emphasize the function of planning in regard to *how* to obtain a given goal in a hetter way. But from about age 7 there is an increase in some other kinds of answers, in particular, the one that planning is useful in order to *perform quickly and save time.* 9-year olds show a *clear* awareness of the benefits of planning for *saving resources,* like money and time, and thus obtaining something that one probably would not have obtained otherwise.

Let us assume that the development in childrens' conceptions of what planning is and of when and how one plans to some extent mirrors a development in when and how children themselves actually plan. Given this assumption there indeed seems to occur *some* development of anticipatory planning in children around 7 years of age.

On the same assumption - and here there is independent evidence that the change in conceptions related by Kreitler and Kreitler is indeed parallelled by a change in actual planning - some other aspects of anticipatory planning, however, become more prominent first in 9-year olds. Here we see *afrequent* application of planning for dealing with *imaginary eventualities,* for instance, planning what to do if a thief comes in at night, if the school is attacked and one is the only defender, and so on. This increase of fantasy themes in planning, as Kreitler and Kreitler suggest, is probably developmentally important in that it may prepare the ground for a further liberation of planning from the sphere of one's immediate interests. The liberation of planning from here-and-now, they claim, becomes very evident in 11-year olds, who consider planning as applicable for domains such as future studies, career, marriage, leaving home, and so on. (And there seems to be a parallel development concerning social planning, that is,. in an extension to consider other agents as agents when making plans, to make plans for others and so on.) (*Ibid.,* pp.216-218.) Furthennore, children around age 9 already have a grasp of both the positive and negative experiential sides of planning, as is reflected in these utterances: "When you do something you planned, you feel good."; "It is boring when you know everything that will happen"; "You enjoy less the action"; "You think only of the plan and not of what there is". *(Ibid.,* p.222.)

The ontogenetic development of anticipatory planning, in my view, is an area for future research and studies, as well as the domain of the ontogenetic development of time representation.

13. CONCLUSIONS AND SUGGESTIONS FOR FURTHER RESEARCH

In this final chapter I will give a brief review of some main ideas and perspectives on planning that have been brought up in this thesis, in particular, those that contribute to an understanding of human planning.

First of all, the planning of action is given a place as a *natura! biological* phenomenon, where the capacity for planning is viewed as being based on functions of perception and of evolutionally earlier forms of behaviour control.[14] That is, the advanced capacity of figuring out, structuring and planning actions that we observe in humans - and which seems to underlie some of the achievements and progress of humankind - is given a certain basis as a biological phenomenon, in spite of the fact that it indeed seems to be a sophisticated capacity which is difficult tograsp.

The evolutionary perspective turned out to be helpful for illuminating the human capability for planning in some of its beneficia! and detrimental aspects.

In particular, some light has been shed on human planning by the characterization and discussion of *anticipatory planning* in contrast to immediate planning. The idea that *time representation* is an important element involved in planning is of course not original, but this thesis has given some suggestions as to a specification of its role.

Furthermore, *the systems perspective,* which means that the capacity for planning is not treated in isolation but as a property of a whole system that perceives, acts, conceives and feels, as well as plans, allowed certain crucial questions to be addressed which are otherwise seldom discussed. From the systems perspective, it is neither natural to view the *making ofplans* as a *seif-evidently adequate* and desirable means for problem solving nor to *equate* 'planning' with 'problem salving'. Instead, I have addressed the issue of the *value* of planning; the value of a particular planning capacity and the value of planning in different situations. By dealing thusly with questions of *why* planning and not just of *how,* the thesis has, I hope, shed some light on planning in general as well as on human planning.

In particular, the suggestion that planning in certain respects is intimately associated with *playing* offers ways of conceiving of the human planning capacity in its adaptive and non-adaptive aspects, and in this concluding chapter I want to dwell a bit more on this. On the positive side, a *planning capacity* - in particular, a capacity for anticipatory planning - as well as a *capacity for playing,* involves a certain *autonomy* and *creativity.* The *playing* individual is *autonomous* in the sense that it is relatively independent of the particular limitations of its present environment. It may, for instance, play around with clay or with cloth or with sticks or whatever he finds, perhaps building something out of this material. As it is not going to be seriously used, the properties of the material are not

[14]This view of planning is *not* obtained simply by using a *wide notion* of planning and plans, where any kind of program or structure that can be read out from behaviour is a *plan* and where any adaptive behaviour is *planned.*

crucial. The individual is *creative* in the sense that it tests things, tries and combines actions in novel ways and improves its own capabilities. It does what it has never done befare, learns about possibilities and discovers new things and novel ways of doing things.

The *planner,* like the playing individual, is *autonomous* in the sense that it undertakes something that is not strictly constrained by its present environment. Planning isa kind of thinking through and playing around intemally with *possible* environments, situations and actions. In particular, *anticipatory planning* means a detachment from both one's present environment and one's present self (one's needs, desires, etc). For instance, when making plans for the summer, I am not constrained by the fact that it is snowing and that I am now cold, or that I have not yet finished my thesis, or that I have no money for a vacation *now.* And in playing around, thinking about new actions and new ways of doing things and in discovering possibilities, the planner can be *creative.* If the plan is acted out, he may indeed have created something in a *double sense;* namely, by creating an action pattem or structure first in an intemal world and then in the extemal world.

These aspects of autonomy and creativity are *evolutionally valuable,* i.e., they increase the inclusive fitness of individuals. They also involve *experiential value,* in that an individual may obtain satisfaction from being autonomous (to have some control, to be independent and not just pushed round by the environment) and creative (to discover things and do novel things and become hetter at <loing things).

On the negative side, I have pointed out that to some extent, however, non-adaptive planning - anticipatory planning in particular - may counteract some of the values of good planning and of play. Starting out from the capacity for playing, which is the most biologically basic and evolutionally prior capacity of the two, planning - anticipatory planning in particular - may, I have argued, come to *replace* play to some extent. *Internal play* replaces *externalplay.* There is *more* playing around with intemal possibilities and *less* with extemal possibilities. The agent may think of future possibilities, plan and have worries and concems for the future *at the expense of* playing around with possibilities that exist in the immediate surrounding; that is, pay more attention to and think more about future possibilities than the present situation. This will often be non-adaptive from an evolutionary perspective as well as from an individual experiential point of view.

On the whole, this thesis has dealt with issues of an encompassing nature, and consequently many proposals are speculative. Its main value may therefore be that it provides hypotheses and perspectives that can be used as a departure for future investigations. In particular, I think of the following topics for further research:

• A more systemarie examination of *anticipatory planning* and *anticipatory behaviour* from an evolutionary perspective.

• More research conceming planning in primates and planning in early humans (for instance, research into the context of its evolution).

• A comparison between the circumstances surrounding planning for early humans and humans in modem society

• Studies of planning, in particular anticipatory planning, in children.

• Studies of the relationships between time representations and planning, in particular anticipatory planning, from a developmental, a cross-cultural and an evolutionary perspective.

• Investigations of *disturbances* in the capacity of planning and in the conception of time in humans.

• A systematic study concerning the value of planning in general, and of anticipatory planning in particular, from different perspectives.

• The development of a simple model for investigating the advantage of an anticipatory planner (or a group of anticipatory planners) over an immediate planner (or a group of immediate planners).

• A comparison between an immediate and an anticipatory planner in various planning situations.

Most of these issues are such that they require, or at least would benefit from, the perspectives of more than one discipline, in particular, psychology, anthropology and philosophy but also AI and neuro-science. In brief, the area is part of cognitive science.

Finally, after having studied planning in theory and in practice during several years, I retain my original fascination for the phenomenon. In particular, the human capacity for anticipatory planning is, I contend, an intriguing and powerful aspect of humans as cognitive and biological systems.

A good plan of action is one that leaves)Ot;\t-4 head and tt,\..-nS into actions and expe..-iences that a..-e frt,\ly of)Ot;\t-4 own design, while yet leaving opportt,\nit fo..- St;\t-4p..-is and imp..-ovisation.

BIBLIOGRAPHY

Agre, P., (1990), 'The role of plans in activity'. In *The Twelfth Annual Conference of the Cognitive Science Society,* 1990, Lawrence Erlbaum, Hillsdale, New Jersey.

Agre, P.E., and D. Chapman, (1987), 'Pengi: An implementation of a theory of activity'. In *Proceedings of National Conference on Artificial Intelligence,* Seattle, WA, 268-272.

Alterman, R., (1988), 'Adaptive planning', *Cognitive Science,* 12, 393-421.

Anderson, J.R., (1983), *The Architecture of Cognition,* Harvard University Press, Cambridge, Mass.

Bach, E., (1981), 'On time, tense, and aspect. An essay in English metaphysics'. In P. Cole (Ed.), *Radical Pragmatics,* New York: Academic Press, pp. 63-81.

Bischof, N., (1980), 'On the phylogeny of human morality', in G.S.Stent (Ed.), *Morality as a Biological Phenomenon,* University of California Press, Berkley, pp. 48-66.

Bischof, N., (1985), *Das Ratsel Odipus,* Piper, Munchen.

Bischof-Kohler, D., (1985), 'Zur Phylogenese menschlicher Motivation'. In L.H. Eckensberger and E.-D. Lantermann (Eds.), *Emotion und Reflexivitat,* Urban & Schwarzenberg, Wien, pp. 3-47.

Boden, M.A., (1983), 'Artificial intelligence and animal psychology', *New Ideas in Psychology,* 1, 11-33.

Boden, M.A., (1977), *Artificial Intelligence and Natural Man,* Harvester Press, Brighton.

Bruce, V., and P.R. Green, (1990), *Visual Perception - Physiology, Psychology and Ecology,* Lawrence Erlbaum, Hove.

Bubot, M.-C., and H. Poucet, (1987), 'Role of the spatial structure in multiple choice problem-solving by golden hamster'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 124-134.

Campbell, D.T., (1966), 'Evolutionary epistemology'. In P.A. Schlipp (Ed.), *The Philosophy of Karl Popper,* Open Court Publishing, La Salle.

Chapuis, N., (1987), 'Detour and shortcut abilities in several species of mammals'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 97-106.

Clark, A., (1986), 'A biological metaphor', *Mind & Language,* 1, 1, 45-63.

Cohen, P.R., and E.A. Feigenbaum (Eds.), (1982), *The Handbook of Artificial Intelligence,* Volume 3, William Kaufmann, Los Altos, Cal.

Craig, W., (1918), 'Appetites and aversions as constituents of instincts', *Biological Bulletin,* 34, 2, 91-107.

Darwin, C., (1872), *The Expression of Emotions in Man and Animal,* Appleton, London.

Dawkins, R., (1976), *The Selfish Gene,* Oxford University Press, Oxford.

De Lisi, R., (1987), 'A cognitive-developmental model of planning'. In S.L. Friedman, E. Kofsky Scholnick, and R.R. Cocking (Eds.), *Blueprints for thinking,* Cambridge University Press, Cambridge, pp. 79-109.

Dennett, D., (1978), *Brainstorms: Philosophical Essays on Mind and Psychology,* Harvester, Brighton, (edition 1981).

Dennett, D., (1983), 'Intentional systems in cognitive ethology: the 'panglossian paradigm' defended', *Behavioral and Brain Sciences,* 6, 343-390.

Dennett, D., (1984a), 'Cognitive Wheels: the Frame Problem of Artificial Intelligence', in C. Hookway (Ed.), *Minds, Machines and Evolution,* Cambridge University Press, Cambridge.

Dennett, D., (1984b), *Elbow Room,* The MIT Press, Cambridge, Mass.

Eibl-Eibesfeld, I., (1978) *Grundriss der vergleichenden Verhaltensforschung,* Piper, München.

Ellen, P., and C. Thinus-Blanc (Eds.), (1987), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht.

Ellen, P., (1987), 'Cognitive mechanisms in animal problem-solving'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 20-38.

Ellis, H.C., and R.R. Hunt, (1983), *Fundamentals of Human Memory and Cognition,* Wm. C. Brown Company Publishers, Dubuque, Iowa

Fabrigoule, C., (1987), 'Study of cognitive processes used by dogs in spatial tasks'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 114-123.

Flavell, J.H., (1977), *Cognitive development,* Prentice-Hall, Englewood Cliffs, NJ.

Fodor, J., (1983), *The Modularity of Mind,* Bradford Books/The MIT Press, Cambridge, Mass.

Freyd, J.J., (1987), 'Dynamic Mental Representations', *Psychological Review,* 94, 4, 427-438.

Friedman, S.L., E. Kofsky Scholnick, and R.R. Cocking (Eds.), (1987), *Blueprints for thinking,* Cambridge University Press, Cambridge.

Gardner, H., (1982), *Developmental Psychology,* Little, Brown and Company, Boston.

Gardner, H., (1983), *Frames of Mind,* Paladin Books, Granada Publishing, London, (edition 1985).

Gardner, H., (1985), *The Mind's New Science,* Basic Books, Inc., New York, (edition 1987).

Gallup, G., (1977) 'Self-recognition in primates: A comparative approach to bidirection properties of consciousness', *American Psychologist,* 32, 329-338.

Geertz, C., (1973), *The Interpretation of Culture,* Basic Books, New York.

Gibson, J.J., (1979), *The Ecological Approach to Visual Perception,* Lawrence Erlbaum, Hillsdale, N.J., (edition 1986).

von Glasersfeld, E., (1976), 'The Development of Language as Purposive Behaviour'. In Harnad *et al.* (Eds.), *Origins and Evolution of Language and Speech. Annals of the New York Academy of Sciences,* Volume 280, pp. 212-226.

Glass, A.L., and K.J. Holyoak, (1986), *Cognition,* Random House, New York.

Goodnow, J.J., (1987), 'Social aspects of planning'. In S.L. Friedman, E. Kofsky Scholnick, and R.R. Cocking (Eds.), *Blueprints for thinking,* Cambridge University Press, Cambridge, pp. 179-204.

Goude, G., (1977), 'Man - a biological being or a technological mistake? A need to decide how to make decisions', Report 4 - 77, Department of Psychology, Gothenburg, Unpublished paper.

Hayes-Roth, B., and F. Hayes-Roth, (1979), 'A cognitive model of planning', *Cognitive Science,* 3, 275-310.

Holst, E. von, (1969), *Zur Verhaltensphysiologie bei Tieren und Menschen. Gesammelte Abhandlungen, Bd. 1 und 2*, Piper, Mtinchen.

Honig, W.K., (1987), 'Local cues and distal arrays in the control of spatial behavior'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 73-88.

Hume, D., (1739), *A Treatise of Human Nature,* (1962 edition).

James, W., (1890), *The Principles of Psychology,* Henry Holt, New York.

Kahneman, D., and A. Tversky, (1972), 'Subjective Probability: A judgment of representativeness', *Cognitive Psychology,* 3, 430-454.

Kahneman, D., and A.Tversky, (1973), 'On the psychology of prediction', *Psychological Review,* **80,** 237-251.

Kahneman, D., and A.Tversky, (1973),' Availability: A heuristic for judging frequency and probability', *Cognitive Psychology,* **5,** 207-232.

Kreitler, S., and H. Kreitler., (1987), 'Plans and planning: their motivational and cognitive antecedents'. In S.L. Friedman, E. Kofsky Scholnick, and **R.R.** Cocking (Eds.), *Blueprints for thinking,* Cambridge University Press, Cambridge, pp. 1107' 178.

Kreitler, S., and H. Kreitler., (1987), 'Conceptions and processes of planning: the developmental perspective'. In S.L. Friedman, E. Kofsky Scholnick, and **R.R.** Cocking (Eds.), *Blueprints for thinking,* Cambridge University Press, Cambridge, pp. 205-272.

Lloyd D., (1987), 'Mental representation from the bottom up', *Synthese,* **70,** 23-78.

Lorenz, K., (1973), *Die Ruckseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens.* Piper, Munchen.

Maynard Smith, J., (1975), *The Theory of Evolution,* Penguin, Harmondsworth.

Maynard Smith, J., (1978), *The Evolution of Sex,* Cambridge University Press, Cambridge.

Marr, D., (1982), *Vision; A Computational Investigation into the Human Representation and Processing of Visual Information,* W.H. Freeman, San Fransisco.

Meacham, J.A., (1984), 'The social basis of intentional action', *Human Development,* **27,** 119-124.

Menzel, E.W., (1973), 'Chimpanzee spatial memory organization', *Science,* 182, 943-945.

Menzel, Jr., E.W., (1987), 'Behavior as a locationist views it'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 55-72.

Miller, O.A., Galanter, E., and K.H. Pribram, (1960), *Plans and the Structure of Behavior,* Holt, Rinehart & Winston, New York.

Miller, O.A., (1956), 'The magical number seven, plus or minus two: Some limits on our capacity for processing information', *Psychological Review,* **63,** 81-97.

Neisser, U., (1976), *Cognition and Reality,* W.H. Freeman, New York.

Newell, A., and H. Simon, (1972), *Human problem solving,* Englewood Cliffs, New York, Prentice-Hall.

Oatley, K., (1978), *Perceptions and Representations,* Methuen & Co, London.

Oppenheimer, L., (1987), 'Cognitive and social variables in the plan of action'. In S.L. Friedman, E. Kofsky Scholnick, and R.R. Cocking (Eds.), *Blueprints for thinking,* Cambridge University Press, Cambridge, pp. 356-392.

Parker, S.T., and K.R. Gibson, (1979), 'A developmental model of the evolution of language and intelligence in early hominids', *Behavioral and Brain Sciences,* 2, 367-407.

Powers, **W.T.,** (1973), *Behavior: The Control of Perception,* Aldine Publishing Company, Chicago.

Premack, D., (1976), *Intelligence in Apes and Man,* Erlbaum, New Jersey.

Premack, D., and G. Woodruff, (1978), 'Chimpanzee problem-solving: a test for comprehension', *Science,* **202,** 532-535.

Premack, D., and G.Woodruff, (1978), 'Does the chimpanzee have a theory of mind?', *Behavioral and Brain Sciences,* **4,** 515-26.

Pugh, G.E., (1977), *The Biological Origin ofHuman Values,* Routledge & Kegan Paul, London.

Riedl, R., (1979), *Biologie der Erkenntnis,* Paul Parey, Berlin.

Rozen, P., (1976), 'The evolution of intelligence and access to the cognitive unconscious'. In *Progress in Psychobiology and Physiological Psychology,* Volume 6, Academic Press, New York.

Runeson, S., (1983) 'Kan man se hur tung en låda ar?', *Forskning och Framsteg,* 5, pp. 3-9.

Ruse, M., (1979), *Sociobiology: Sense or Nonsense,* Reidel, Dordrecht.

Sacerdoti, E.D., (1977), *A Structurefor Plans andBehavior,* Elsevier, Amsterdam.

Sachsse, H., (1971), *Einfuhrung in die Kybernetik,* Friedr. Vierweg & Sohn, Braunschweig.

Savage-Rumbaugh, E.S., Rumbaugh, D.M., and S. Boysen, (1978), 'Linguistically mediated tool use, and exchange by Chimpanzees *(Pan troglodytes), Behavioral and Brain Sciences,* **4,** 539-554.

Schachter, S., and J. Singer, (1962), 'Cognitive, social and physiological determinants of emotional states', *Psychological Review* , **69,** 379-99.

Schank, R.C., and R.P. Abelson, (1977), *Scripts, Plans, Goals, and Understanding,* Lawrence Erlbaum, Hillsdale, NJ.

Seyfarth, R.M., Cheney, D.L., and P. Marler, (1980), 'Vervet monkey alarm calls: semantic communication in a free-ranging primate', *Animal Behaviour,* 28, 1070-1094.

Shephard, R.N., (1984), 'Ecological constraints on internal representation: resonant kinematics of perceiving, imagining, thinking, and dreaming', *Psychological Review,* 91, 4, 417-447.

Sherry, D.F., (1987), 'Spatial memory in food-storing birds'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 305-322.

Simon, H., (1969), *The Sciences of the Artificial.,* MIT Press, Cambridge, Mass.

Sloman, A., (1987), 'Motives, mechanisms, and emotions', *Cognition and Emotion,* **1,** 3, 217-233.

Stent, G.S. (Ed.), (1980), *Morality as a Biological Phenomenon,* University of California Press, Berkley.

Suchman, L.A., (1987), *Plans and Situated Actions: The Problem of Human-Machine Communication,* Cambridge University Press, Cambridge.

Thinus-Blanc, C., (1987), 'The cognitive map concept and its consequences'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 1-19.

Tinbergen, N., (1951), *The Study of Instinct,* Oxford University Press, London.

Tulving, E., (1972), 'Episodic and semantic memory'. In E. Tulving, and W. Donaldson (Eds.), *Organization of memory,* Academic Press, New York.

von Uexkiill, J., (1970), *Streifzuge durch die Umwelten von Tieren und Menschen, Bedeutungslehre,* Fischer Wissenschaft, Frankfurt am Main.

Vauclair, J., (1987), 'A comparative approach to cognitive mapping'. In P. Ellen, and C.Thinus-Blanc (Eds.), *Cognitive Processes and Spatial Orientation in Animal and Man,* Volume I, Martinus Nijhoff Publishers, Dordrecht, pp. 89-96.

Vollmer, G., (1975), *Evolutionare Erkenntnistheorie,* Hirzel, Stuttgart.

Whiten, A., and R.W. Byme, (1988), 'Tacticaldeception in primates', *Behavioral and Brain Sciences,* **11,** 2, 233-244.

Whorf, B.L., (1956), *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whord,* ed. John B. Carroll. MIT Press, Cambridge, Mass.

Wilensky, R., (1981), 'Meta-planning: Representing and using knowledge about planning in problem solving and natural language understanding, *Cognitive Science,* 5, 197-233.

Wuketits, F.M., (1981), *Biologie und Kausalitat,* Paul Parey, Berlin.

tralen
Lunas 'Universitet 1991