Unsupervised navigation using an economy principle

Lawrence Warnett & Brendan McGonigle[*]
Laboratory for Cognitive Neuroscience & Intelligent Systems
University of Edinburgh
Edinburgh, Scotland
ejua48@holyrood.ed.ac.uk

Abstract

We describe robot navigation learning based on self-selection of privileged vectors through the environment in accordance with an in built economy metric. This provides the opportunity both for progressive behavioural adaptation, and adaptive derivations, leading, through situated activity, to "representations" of the environment which are both economically attained and inherently meaningful to the agent.

1. Introduction

A central objective of our programme at Edinburgh has been to help specify core conditions of growth in cognitive competencies and cognitive regulation in complex intelligent systems over the course of their evolution and development (McGonigle and Chalmers, 1996, McGonigle and Chalmers, 1997, McGonigle and Chalmers, 2002, for further information). An adaptive, epigenetic stance, this has led in turn to the design of robots with functional architectures which support different types of learning in an interdependent way.

2. An epigenetic landscape for robots based on both design and learning processes.

The design features do not mean that behaviours have been installed. Learning processes play a significant role. Rather than utilise a general purpose learning process, we adopted a "horses for courses" evolutionary approach which suggests different types of learning mechanisms within different layers of adaptation. Thus our learning devices tailored first to basic reactive layers, have been based on instinct rules" (Nemhzow et al., 1993). Our second form of learning was based on anticipation of clutter using route learning which geared a hearing robot's speed according to the route segments logged by the robot on various trials as causing the highest levels of perturbation from objects in its path (Donnett and McGonigle, 1991). Enabling anticipation of hazard, such learning affords much more economic journeys to the goal (in this case, a sound source). Later we applied principles from operant conditioning in the form of external tuition based on 'shaping'. This produced dramatically fast learning of a free ranging but supervised robot (Nehmzow and McGonigle, 1995 see http://bion.psy.ed.ac.uk). And this has been followed by reinforcement learning applications inspired by Humphrys (Humphrys, 1996a, Humphrys, 1996b, Humphrys, 1997) designed to cope with multitasking where we have considered learning by internal competition between numerically-scored alternatives (McGonigle, 2002, for further details).

Whilst effective in part, these approaches have all had their problems. In reinforcement learning, shaping gives a learning algorithm 'hints' by modifying the reward function but achieves tractability by frequently changing the problems in unanticipated ways that cause poor solutions to be learnt. Furthermore, naïve algorithms often scale exponentially in the number of state variables, and are thus frequently impractical.

3. Unsupervised learning

In the 'learning to navigate'example we report here, we have implemented a type of self-organised learning as revealed in our primate cognitive research and motivated by a cognitive economy policy (McGonigle and Chalmers, 2002, McGonigle and Chalmers, 1997). Based on search tasks which demand high levels of serial, executive control in the sequencing of large numbers of icons on touch sensitive screens, we have explored new types of learning based on the subjects' self selection of those procedures (for example classification principles and chunking) which enable the most effective forms of executive performance with the least investment in cognitive resource (see http://bion.psy.ed.ac.uk for examples). Strategic, not merely dispositional, the sorts of procedures selected by the subject vary with the task demands.

Akin to animals seeking to optimise foraging to procure the greatest return for the least investment in energy expended, we have engineered a new form of navigation, having the robot select from magnetic compass information the longest axis in its work space it can move along with the least perturbation caused by obstacles.

Logical (Functional) Hierarchy

Active vision

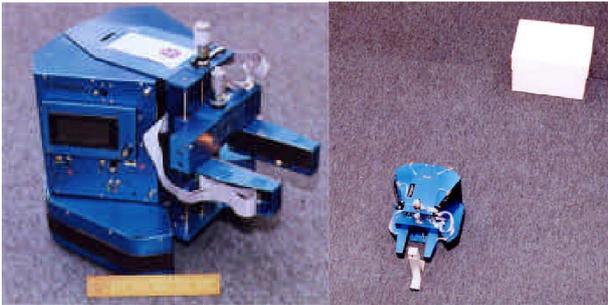Place Learning: navigation

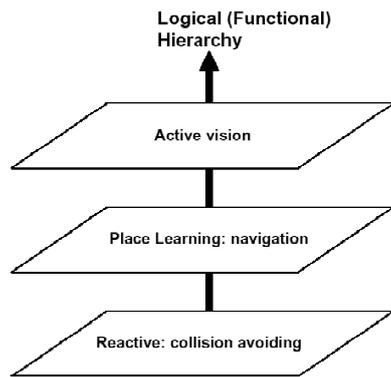Reactive: collision avoiding

Figure 1. Design. Three layers of competences have been installed as a design feature of each agent based on a logical hierarchy as described by McGonigle (1990). In common with robots designed by Brooks at MIT (Brooks, 1991, for example), the first layer is purely reactive, involving tightly coupled 'first strike' behaviours designed to enable the robot avoid collision. Crucially, however, this has only been a first stage; in the second layer of the hierarchy, competences based on sound sources (Donnett and McGonigle, 1991), then on a light compass and dead-reckoning (Nehmzow and McGonigle, 1995), and more recently a magnetic compass, have been engineered, enabling the robot to identify locations which are of significance in their workplace (Fischer et al., 2000). With navigation in place, an active vision- based form of object identification has been achieved, scaffolded by the precursor achievements enabled by the lower layers of the hierarchy.

Acting in real-time and using relatively low-level, and highly error laden sensory-motor information combined with a primitive set of state variables, we have sought to develop a system which could learn to self-select optimal trajectories in our laboratory environment guided only by a resource minimisation constraint. In the case of the test niche as illustrated above, only trajectories along the long corridor will qualify as those which enable the longest travel with the least perturbation.

4. The hierarchical architecture underlying self-organised navigational learning

The Nomad has access to data structures and programs that are arranged hierarchically from atomic units of action (for example move, read-sensor, talk) many of which come pre-installed, through simple sequences of actions which combine both movement and sensing atoms (an everyday example from robotics is move-forward-without-crashing), and fully to combinations of actions which form more meaningful, or rather non-trivial, sequences which we designate behaviours (such as navigate-to-position, align, learn-route, retrace-route). The top-most level of the hierarchy consists of tasks e.g. groups of behaviours analogous to those in natural complex systems such as learning and retracing routes through the niche. This architecture is broadly object-oriented, with behaviours comprising complex data structures which retain information concerning: the number of calls and successful instantiations; initialisation constraints (both internal and external criteria); instantiation & execution; runtime constraints (success and failure conditions); lists of implementable recovery procedures (which also retain data concerning the number of times they have been called and their success rate); and finally procedures for modifying a subset of internal state variables.

5. Using memory (both "long" and "short" term) to guide the developmental trajectory

The Nomad's control architecture features two sets of global variables which reference internal and external states of affairs. Within a single run, in the \short term", these state arrays store information concerning the Nomad's current position in space and time, as well as in task-space and behaviour-space. Not only do these records of previously encountered conditions allow the system to benefit from past experience, but they also facilitate the use of the economy metric used to select promising vectors. Consequently at the end of a run the Nomad stores all information relating to behaviours, tasks, and the state arrays. At any point in time information from earlier in the current run or from any recorded point in the past can be re-loaded. Thus the capacity for true development is achieved and the Nomad freed from the short-termism characteristic of many artificial agents!

6. The trial and error stage

Initially the Nomad is positioned relatively near one end of the long axis of the environment. The pre-existing alignment behaviour positions the Nomad at its niche-engineered home" location with an accuracy of ±5cms. Alignment to features of the environment, unique within the niche, allows the system to orient itself at start-up without the need for resource demanding visual or other sensory-fusion place recognition algorithms and, crucially, ensures that the Nomad begins each run from a near identical point in space allowing for learning across runs. Our learning algorithm begins with varying the initial compass reading by ± 30 and 60 degrees (we start with deliberately coarse gradations) to generate 5 potential headings (vectors). The Nomad then travels forward along each of these vectors in a random order until either an obstacle is detected ahead, or the discrepancy between the current compass reading and the initial bearing reaches 15 degrees 3, whereupon it recenters, essentially reversing back to the initial position.
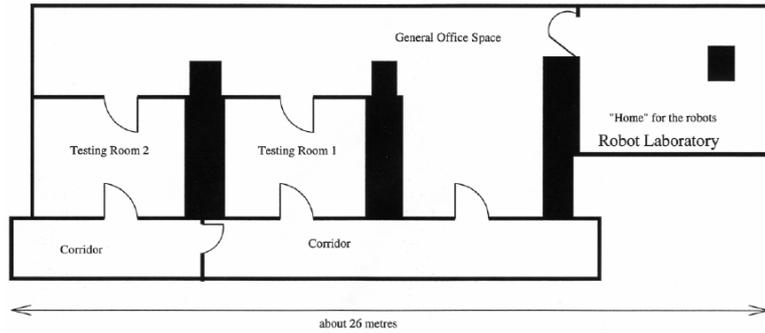
Figure 2. Robot Self Regulation: The Nomad and test niche The robot used for this implementation is a Nomad 200 (see Figure 1), supplied by Nomadic Technologies Inc., California. Its sensors consist of sixteen infrared detectors, and sixteen sensors, located around the periphery of the chassis, along with base mounted odometer and a magnetic compass (KVH C100 Compass Engine). Remote control of the nomad is achieved via a radio Ethernet card, allowing control from any one of a number of machines situated throughout the niche. The Nomad is programmed in a mixture of the C/C++ and Perl programming languages, running under the Red Hat Linux operating system. Each language provides its own benefits .C++ is used o communicate with the onboard robot control system, and is particularly useful in the control of time-critical processes. The strength of Perl running under Linux lies in the control of concurrent processes, and the dynamic construction of both programs and data structures

A log is maintained of odometric information, compass discrepancy, and time travelled in each direction. Once the set of potential vectors has been exhausted the most promising one is determined. If there is a vector where compass error has exceeded 15 degrees (in such cases the distance travelled tends to be relatively great due to the additive error in the magnetic compass with distance) it becomes the chosen vector for the current sector, if more than one vector meets this criterion, that which admits the longest run for the least error is chosen. If no vectors meet the success criterion, then the one along which there has been the longest run without interrupt is chosen. Hereupon the Nomad travels in the chosen direction until the sector boundary is reached (determined by the odometry, compass error and time logs). The process is then iterated indefinitely, while a record is stored of which sector (or, physical area) the Nomad is currently positioned within. This procedure results in movement along free axes of the niche, which is broken down by selfdetermined internal criteria into discrete physical spaces. Returning to base invokes a retrace-route procedure implemented by a process of dead reckoning where possible. Indeed the structure of the Nomad's niche (long corridors, with separate rooms, and dog-leg corners), means that the strategy of steering more or less directly toward home, in the absence of planning and prospective control, often fails, thereby furnishing the system with a rich error space in which to test recovery procedures. Although, at this time, only preliminary research has been carried out in this area with the Nomad, preliminary findings are promising. Our initial investigations focussing on two simple methods of recovering from navigational error (wall following, and reiteratively creating a temporary goal location midway between current and target spatial positions) indicate that the concept of a stack of error recovery procedures based on abductive 'explanation' procedures, is workable in principle.
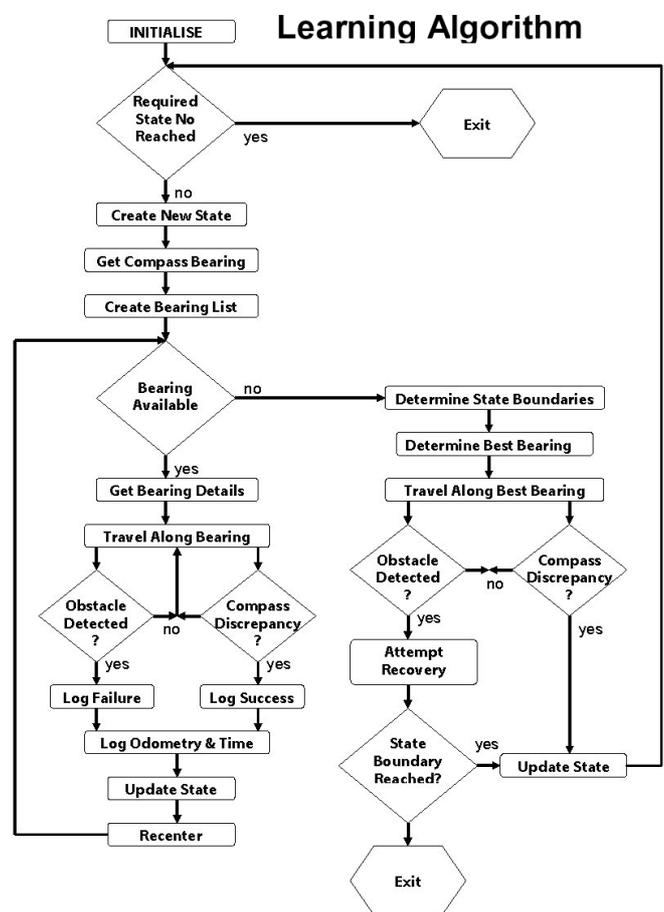

Figure 3. Learning Algorithm

References

Brooks, R. A. (1991). Intelligence without representation. Artificial Intelligence, 47:139-159.

Bryson, J. and McGonigle, B. O. (1999). Agent architecture as object-oriented design. In Singh, M. P., Rao, A. S., and Wooldridge, M. J., (Eds.), The fourth international workshop on agent theories, architechtures and languages (ATAL 97, Providence. Springer-Verlag.

Donnett, J. and McGonigle, B. O. (1991). Evolving speed control in mobile robots: from blindnessto kinetic vision. In Proceedings of the Vision Interface Conference, Calgary.

Duckett, T. and Nehmzow, U. (1999a). Knowing your place in real world environments. In Proceedings EUROBOT'99, 3rd European workshop on advanced mobile robots. Zurich, Switzerland, September 6-8, 1999.

Duckett, T. and Nehmzow, U. (1999b). Self-localisation and autonomous navigation by a mobile robot. Technical report UMCS-99-3-1,Department of Computer Science, University of Manchester, England. also in Proceedings TIMR 99 "Towards intelligent mobile robots" Bristol, England.

Fischer, K., Zheng, Y., McGonigle, B. O., and Warnett, L. (2000). Learning and developing together: dynamic construction of human and robot knowledge. In McClelland, J. L. and Pentland, A. P., (Eds.), NSF DARPA Workshop on Development and Learning, pages 50-59.

Humphrys, M. (1996a). Action selection in a hypothetical house robot: using those RL numbers. In Anderson, P. G. and Warwick, K., (Eds.),Proceedings of the First international ICSC symposia on intelligent industrial automation (IIA-96) and soft computing, pages 216-222, Reading, England. ICSC Academic Press. March 26-28.

Humphrys, M. (1996b). Action selection methods using reinforcement learning. In Maes, P., Mataric, M., Meyer, J.-A., Pollack, J., and Wilson, S. W., (Eds.), From animals to animats 4: Proceedings of the fourth international conference on the simulation of adaptive behaviour, pages 135-144, Cambridge, MA, USA. MIT Press.

Humphrys, M. (1997). Action selection methods using reinforcement learning. PhD thesis, University of Cambridge, Cambridge, England.

Lu, F. and Milios, E. (1997). Globally consistent range scan alignment for environment mapping. Autonomous Robots, 4:333-349.

McGonigle, B. O. (1991a). Incrementing intelligent systems by design. In Meyer, J. and Wilson, S., (Eds.), From Animals to Animats, Proceedings of the First International Conference on the simulation of adaptive behaviour. MIT Press.,London, England.

McGonigle, B. O. (1991b). Incrementing intelligent systems by design. In Meyer, J. A. and Wilson, S. W., (Eds.), From animals to animats: Proceedings of the First international conference on the simulation of adaptive behaviour. MIT Press, London, England.

McGonigle, B. O. (2002). Getting autonomous robots to control themselves. In Holland, O. and McFarland, D., (Eds.), Articial ethology. Oxford University Press, Oxford, England. Invited contribution. International symposium,Lanzarote, 1998, sponsored by OUP.

McGonigle, B. O. and Chalmers, M. (1996). The ontology of order. In Smith, L., (Ed.), Critical readings on Piaget, pages 279-311. Routledge, London, England.

McGonigle, B. O. and Chalmers, M. (1997). Rationality as optimised cognitive self-regulation. In Oaksford, M. and Chater, N., (Eds.), Rational Models of Cognition. OUP, London.

McGonigle, B. O. and Chalmers, M. (2002). Animal cognition and sequential behaviour: behavioural, biological and computational perspectives, chapter 12. The growth of cognitive structure in monkeys and men. Kluwer Academic Publishers.

Nehmzow, U. (1995). Animal and robot navigation.Robotics and Autonomous Systems, 15(1-2):71-81.

Nehmzow, U. and McGonigle, B. O. (1993). Navigation mobiler roboter mittels differential lichtkompass. Hybride und intergrierte Ansatzezur Raumrepraentation und uhre Andewung.Berlin. German journal for Artificial Intelligence, "KI". ISBN 0933-1875.

Nehmzow, U. and McGonigle, B. O. (1995). Achieving rapid adaptations in robots by means of external tuition. In Cliff, D., Husbands, P., Arcady-Meyer, J., and Wilson, S. W., (Eds.), From Animals to Animats 3. Proceedings of the Third International Conference on Simulation of Adaptive Behaviour. MIT Press, Cambridge, MA, USA.

Nemhzow, U., Smithers, T., and McGonigle, B. O.(1993). Increasing behavioural repertoire in a mobile robot. In Meyer, J., Roitblat, H.,and Wilson, S. W., (Eds.), From Animals to Animats 2. Proceedings of the Second International Conference on Simulation of Adaptive Behaviour. MIT Press, Cambridge, Massachusetts;London, England.

Terrace, H. S. and McGonigle, B. O. (1994). Memory and representation of serial order by children, monkeys and pigeons. Current Direction in Psychological Science, 3(6):180-185.

Warnett, L. (2000). Intelligent systems - towards a new synthetic agenda. PhD thesis, Department of Psychology, University of Edinburgh, Edinburgh, Scotland.

Weiss, G. and von Puttkamer, E. (1995). A map based on laserscans without geometric interpretation. In Intelligent autonomous systems-4(IAS-4), pages 403-407, Karlsruhe, Germany.

Yamauchi, B. and Langley, P. (1997). Place recognition in dynamic environments. Journal of Robotics Systems, 14(2):107-120. Special issue on mobile robots.