In Berthouze, L., Kozima, H., Prince, C. G., Sandini, G., Stojanov, G., Metta, G., and Balkenius, C. (Eds.)
Proceedings of the Fourth International Workshop on Epigenetic Robotics
Lund University Cognitive Studies, 117, ISBN 91-974741-3-4

Sharing Meaning with Machines

Claire D'Este
School of Computer Science and Engineering
University of New South Wales
Sydney, NSW 2052 Australia
claired@cse.unsw.edu.au

Abstract

Communication can be described as the act of sharing meaning, so if humans and machines are to communicate, how is this to be achieved between such different creatures? This paper examines what else the communicators need to share in order to share meaning, including perception, categorisation, attention, sociability and consciousness. It compares and takes inspiration from communications with others with different perception and categorisation, including the deaf-blind, the autistic and animals.

1. Introduction

As machines become larger parts of our everyday lives and begin to share our environment with us, it has become increasingly desirable to be able to communicate with them in a natural, open-ended way. So if communication is the act of sharing meaning (Tubbs and Moss, 1994), then how can this be achieved between creatures as different as humans and machines?

There is little doubt that a machine can process sensory information, and use some mathematical technique to sort and generalise this into higher-level categories. (Steels and Kaplan, 2002) have demonstrated successful bootstrapping of shared meaning between robotic agents. However, the agents created their own language grounded in their specific and identical perceptual and categorisation systems. We are still left with the issue of how we make machines learn *our* language while made of different stuff.

This problem is not limited to humans and machines. We also want to communicate with humans with extreme sensory or mental impairment, and with animals, despite dissimilar sensory and nervous systems. This paper will examine how shared meaning is established with these very different conversation partners, and how it highlights the necessary and sufficient abilities for a communicating machine.

2. Sharing Perception

To give us something to talk about, we need to share an experience. The focus for machines has been on the visual experience, as sight appears to be the primary sense of humans. However, many animals provide evidence that sight is not necessary for conceptualisation, and sensory-impaired children, although they require a more active and complex teaching process, are eventually able to communicate fluently using the same language.

Teachers of deaf-blind children begin with a limited set of words and do not give the child more until they learn them in various contexts (Witt, 2004). Machines are also usually taught in a restricted world with a restricted vocabulary (Steels and Kaplan, 2001). Deaf-blind children can be taught to speak by reading the vibrations on people's lips and feeling their own vibrations as they speak. But small children are taught gesticulation and finger spelling first to learn the meaning of the word before they are permitted to learn how it is spoken. Without previous knowledge of the word the child may have to repeat the utterance up to ten thousand times before understanding it. Deaf-blind children have the disadvantage of not being "bathed in words" from the moment of their birth (Colligan, 1962). This makes the words far more distant from their meanings than with actions that are close to those actions they perform on objects everyday, or shapes that are similar to the object. The gestures are then easily mapped to spoken words.

If touch alone can ground language, it is unfortunate that the technology is still limited and expensive, as it would be interesting to clothe a robot in an artificial skin and guide their sensing. Vision alone is perhaps too passive.

3. Sharing Concepts

Once the body has sensed things in its environment the next step is to categorise into concepts. Meanings can be defined as being concepts alone (Horwich, 2003). Unfortunately, it is not possible to directly observe the conceptual systems of humans, or to compare and copy them. Philosophers, psychologists and neurologists are yet to agree on how humans store their experience, as pictures in the head, attribute and value pairs, causal indexes, or abstract symbols (Peirce, 1897). So unfortunately, we cannot copy directly from us and be sure the representation is the same. We can only make inferences from the usage of a concept, for example if you run from a tiger we can infer that it's fangs, teeth and ferocity are bound up in your concept of it.

When we use language, we are learning to "classify the world in a shared and modifiable way" (Harnad, 1996). This suggests that this classification for language is not necessarily the same as our personal and subjective categorisation. This can be biologically determined. The term 'Umwelt' is used to describe the species-specific objective world of a creature (Deely, 2001), or in our case of a machine. The physical needs of different animals cause them to concentrate on different parts of their environment through different sensory capabilities. To establish shared meaning those involved need a sufficient overlap of Umwelt so they discuss shared experiences. For example, humans and dogs are able to communicate about items that are part of both of our Umwelten, such as squeezy toys and walks, but you may find it more difficult to discuss your investment portfolio, or the subtleties of urine scent discrimination. So it is perhaps acceptable to begin by speaking of things that machines find easy to notice, such as brightly coloured objects that are easily segmented (which invariably happens in robotic experimentation), as the first words taught to apes are invariably things like 'banana'.

Even the concepts of two different humans with fully functioning sensory and cognitive apparatus can be entirely different depending on their experience. Someone who has seen a platypus lay eggs will quickly update the 'bears live young' part of their concept of what a mammal is. If their conversation partner is unaware of this phenomenon then they will still be able to talk far and wide about mammals as long as it is never necessary to make this distinction. It may be said that they do not truly understand one another, as they are not technically talking about *exactly* the same thing, but can it be said that the second person understands nothing the first is saying? He knows enough to carry out a meaningful exchange.

So it is possible that we may have machines whose concept of concepts is incomplete, but can still converse with us meaningfully. As children we create concepts coarsely until proven otherwise. All fluffy things are first 'dog', then 'dog' and 'kitty', and as we experience more, we create more concepts and discriminations. Our ability to continuously modify and update our concepts as we have new experiences makes us an open-ended solution for meaning. Although the concepts are not immediately correct in the shared use of them (a tiger is not a kitty), it does not mean that these first utterances are without meaning.

4. Sharing Attention

Another perspective is that the personal perceptive and conceptual experience does not need to be shared, as long as you can establish joint attention. If we can be sure that we are both talking about the same thing, it is not important exactly how the distinctions are being made. Most autistic children are unable to establish shared visual attention and this seriously impedes their verbal and nonverbal communication (Baron-Cohen, 1995). This can lead to production competence without obvious meaning. For example, in answer to the question 'How are you?' they might answer 'How are you?' (Lovaas, 1977).

Pepperberg (1999) found her parrots were unable to determine what the word was about if the two humans involved in the model-rival technique (explained below) did not share visual attention on the object. Also, in Steels and Kaplan's (2001) robotic experiments, they found communication failure occurred most often because the object of discussion was simply not in the robots line of sight. When the input vision was limited to examples that humans were able to categorise, the success rate improved dramatically.

We need to be sure the correct association is being made between symbol and sensation, and draw the attention of the machine to the object of interest. 'Normal' children mainly use innate, or early-learned gaze-following abilities and teachers of blind children physically guide their hands. Therefore, it seems mechanisms are required to steer the sensing of the robot and gaze following has been developed in robotic systems (Kozima and Ito, 1997) (Breazeal, 2002). But the problem autistic children have with sharing attention is not with determining what someone is looking at, but following the gaze of another to determine the object of their interest. For example, although they can answer correctly a question such as "What is Peter looking at, the blue ball or the red ball?" they answer randomly questions such as "What does Peter *want*, the blue ball or the red ball?" (Lovaas 1977). So in addition, the ability to predict the beliefs and desires of others, or a 'theory of mind', seems necessary.

5. Sharing a Mind

Severely autistic children seem unable to view other humans as intentional beings like themselves, with beliefs and states of mind (Baron-Cohen, 1995). But, if you can view others as being like yourself, then you can imitate them by mapping their actions onto your own possible movements and sounds. Autistic children are taught to imitate non-verbal behaviour before the teachers will attempt to build speech (Lovaas, 1977). Pepperberg's (1999) grey parrot was able to learn much more quickly and flexibly than previous classical conditioning attempts through a technique called model-rival learning. Rather than learning the objects and associated words directly, the learner observes someone else learning the task and can then imitate them.

In Savage-Rumbaugh's (1994) experiments involving pygmy chimpanzees who typed symbols on a keyboard, it was the chimp's child that learned best from observing her mother being taught. Kanzi was eventually able to comprehend complex and novel sentences such as, "Kanzi, if you give Austin your monster mask, I'll let you have some of Austin's cereal". The evidence that he understood the meaning is that he gave Austin the mask and pointed at his cereal. Kanzi was also able to demonstrate his comprehension without expecting immediate reward. Pepperberg and Savage-Rumbaugh both found that consistent reinforcement of the meaning of a word was more effective than a simple food reward on correct production. If the pupil says 'hat' on seeing a hat, it was better to let them touch the hat, then give them a treat.

An early stage in teaching children with deafblindness is to demonstrate to them that their actions can specifically affect others. A machine must also realise this, and make the appropriate predictions about the outcomes of it's behaviour. Only then can we make the correct associations between what was said and what happened.

Seeing others as intentional beings also makes them more interesting. Other creatures can be a source of information, unpredictability and mutual understanding. Pepperberg and Savage-Rumbaugh's teaching methods rely heavily on the subject enjoying social interactions. The parrots did not learn nearly as effectively when they were shown videos of the objects without human involvement (Pepperberg, 1999). Therefore a degree of sociability may be important.

6. Sharing Sociability

It seems there is a bias in us for paying more attention to human sounds, shapes and movements, and when this malfunctions, as in autism, communication can be limited and often non-existent. Severely autistic children appear to view other humans simply as objects. This clearly makes engaging the child in interaction extremely difficult, as the strange noises coming from the bags of skin are not regarded with any more interest than other ambient sights or sounds. This highlights the need for the machine not only to be able to share attention with the human, but also to find interacting interesting. This can be as simple as it concentrating on human skin colour and noises from the frequency spectrum of human speech (Breazeal, 2002), but this is by no means an open-ended solution.

Sociability also involves the gestures, facial expressions and other sounds that make up a large part of human and animal communication. To encompass this additional information there has been recent interest in equipping embodied agents with similar social skills (Breazeal, 2002). Sociable machines hope to make the interaction more natural for humans and allow us to use our highly developed abilities for interpreting and manipulating others. Humans unconsciously project intention onto autonomous agents (Reeves and Nass, 1996), and will even manipulate their utterances to achieve a more satisfying interaction. It may be possible to exploit these tendencies for the benefit of both.

7. Sharing Consciousness

(Wenburg and Wilmont, 1973) define communication as any attempt to achieve understanding. But the possibility of an understanding machine is still widely disputed. (Harnad, 1990) has made clear that the only understanding that currently takes place in our computer programs, is when a human mind interprets the information. A machine manipulates abstract symbols purely by their shape, and attributes no meaning whatsoever to them. The feeling is that input and output behaviour is not enough to prove the machine understands (Searle, 1980). In the case of a machine, translating my email into Chinese, does not mean it understood it in English.

Human-animal communication research suffers from similar criticism. The learning of correct reactions for sounds, pictures or gestures can be explained by classical conditioning and does not necessarily demonstrate the possession of meaning or understanding. Early attempts used food rewards as the primary motivation for the animal. But (Savage-Rumbaugh and Boysen, 1980) admits that chimps can easily acquire skills that appear like language on the surface, without demonstrating the use of symbols or other important interaction behaviour. "Knowing how to use the symbol 'banana' as a way of getting someone to give you a banana is not equivalent to knowing that 'banana' represents a banana. p. 67" (Savage-Rumbaugh, 1994) Similarly, children with deafblindness can be taught segmented and discrete skills that may help them get what they want in certain situations without shared meaning. But there communication development is stuck at a basic stage and they will not learn to communicate fluently and flexibly.

The quantitative and objective aspects of meaning may not be enough for open-ended conversation with humans. It may be necessary to represent the qualitative and subjective facets of concepts. Without feeling the ineffable feelings behind things, it may be impossible to truly understand the meaning of them (Jakab, 2000), and no formal description will be the same as true experience. This knowing-from-feeling is known as qualia and is believed to be only possible in conscious beings.

If we need to be conscious in order to understand, then perhaps this is something we need to simulate in our machines. Not everyone would agree that consciousness *can* be simulated, or attribute consciousness to animals. But if the feelings are subjective then this aspect of meaning is not necessarily shared, even between conscious beings. Therefore, as it is only an internal 'revelation' it may not be necessary for establishing meaning between conversational partners.

Conclusion

In order to initiate communication it seems several abilities need to be in place, some of which are reliant on each other. From the experience of teachers of animals and impaired children, it seems that the most important aspect is an interest in the teacher over other stimuli. This is necessary to be able to guide their attention and to encourage imitation. Without this present in our machines we cannot begin to provide feedback on what we want the machine to learn, which enables us to enforce a certain type of categorisation (however it is represented) that is shared by speakers of the language.

Acknowledgements

Thankyou to Kylie Witt, Deafblind Services Coordinator for the Royal Institute for Deaf & Blind Children Australia.

References

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* MIT Press, Cambridge, MA.

Breazeal, C. (2002). *Designing Sociable Robots.* MIT Press, Cambridge, MA.

Colligan, J. (1962). Teaching deaf-blind children. *Seminar of the Royal National Institute for the Blind*, Shrewsbury, England.

Deely, J. (2001). Umwelt. *Semiotica 134*, 1:125–135.

Harnad, S. (1990). The symbol grounding problem. *Semiotica 134*, 42:335–346.

Harnad, S. (1996). The origin of words: A psychophysical hypothesis. In Durham, W. and Velichkovsky, B., (Eds.), *Communicating Meaning: Evolution and Development of Language.* Erlbaum.

Horwich, P. (2003). From meaning. In Richard, M., (Ed.), *Meaning.* Blackwell, Malden, MA.

Jakab, Z. (2000). Ineffability of qualia: A straightforward naturalistic explanation. *Consciousness and Cognition*, 9:329–351.

Kozima, H. and Ito, A. (1997). The role of shared-attention in human-computer conversation. In *International Conference of Research on Computational Linguistics (ROCLING-97)*, pages 224–228, Taiwan.

Lovaas, O. (1977). *The Autistic Child: Language Development through Behaviour Modification.* Irvington Publishers Inc, New York.

Peirce, C. (1897). Logic as semiotic. In Buchler, J., (Ed.), *Philosophical Writings of Peirce (1955).* Dover, New York.

Pepperberg, I. (1999). *The Alex studies: cognitive and communicative abilities of grey parrots.* Harvard University Press, Cambridge.

Reeves, B. and Nass, C. (1996). *The Media Equation.* Cambridge University Press, New York.

Savage-Rumbaugh, S. Rumbaugh, D. and Boysen, S. (1980). Do apes use language? *American Scientist*, 68:49–61.

Savage-Rumbaugh, S. (1994). *Kanzi: The ape on the brink of the human mind.* John Wiley & Sons, New York.

Searle, J. (1980). Minds, brains and programs. *The Behavioural and Brain Sciences*, 3:417–424.

Steels, L. and Kaplan, F. (2001). Aibo's first words: the social learning of language and meaning. *Evolution of Communication*, 4.

Steels, L. and Kaplan, F. (2002). Bootstrapping grounded word semantics. *Linguistic evolution through language acquisition: formal and computational models.*

Tubbs, S. and Moss, S. (1994). *Human Communication.* McGraw-Hill, New York.

Wenburg, J. and Wilmont, W. (1973). *The personal communication process.* John Wiley & Sons, New York.

Witt, K. (2004). *Personal communication.*