Seminar Notes On ‘Cognition and Robotics’.

Abstract: ‘Artificial intelligence’ used to be about building and using algorithm based computers to solve problems normally regarded as requiring human intelligence. The influence of neuroscience has led to a new breed of ‘computational intelligence’ that exploits the complex dynamics of ‘neural nets’ and increases our understanding of the term ‘autonomous agent’.
It has also raised the possibility that we might be able to understand more about how the human brain works and ledto a greater ability to build robots that learn to adapt to novel environments. Igor Aleksander stressed the importance of depiction and planning in human learning and its relevance in understanding pathological conditions such as Parkinson’s disease. Inman Harvey described how ‘artificial evolution’ could give us the next generation of intelligent robots.

Director of Complexity Research : Eve Mitleton- Kelly
London School of Economics
Houghton Street
London WC2A 2A

Presenters : Igor Aleksander, Imperial College
Inman Harvey, Sussex University

Compiled For The L.S.E.
by Geoffrey J.C. Higgs 30/5/01

Introduction

The use of computers in solving problems normally regarded as requiring human intelligence used to be known as ‘artificial intelligence’ (AI). Such computers used linear logic; ‘straight line’ sequences of instructions known as ‘algorithms’. The computer known as ‘Deep Blue’, for example, was built by Claude Shannon to play chess and beat the Grand Master, Kasparov by using a ‘pruning out’ algorithm which looked at every possible move over the next ten steps of the game and then chose the best. But the working of ‘Deep Blue’ is unlike what goes on in our heads and the biological neurosciences are beginning to have an impact on our ability to built machines that deal with true novelty. Human brains appear to exploit complexity and all that that means in terms of the new science. A new generation of machines, involving a very different way of working, is necessary if we are to really understand and build ‘autonomous agents’ as robots. On the other hand, understanding the way we as humans cope with the world is an aspect of psychology which is important in treating pathological conditions such as Parkinson’s disease.

Features of Human Perception and Conception

A simple way of understanding ‘seeing’ in the past was that it involved assigning ‘labels’ to pictures coming into our heads via the light waves. We can build robots that operate this way, but it’s really not the way we ‘see’ at all, and it does not account for the enormous amount of activity that goes on in our heads. The eye contains a central part called the fovea and one that surrounds it called the perifovea. But it is a small structure at the back of the eyeball called the ‘superior colliculus’ which moves the eye around and the fovea ‘sees’ differently from the perifovea. In observation there is a great deal of movement of the eye and interestingly a patient with Parkinson’s disease moves his or her eye in a very different way to a normal person. Visual ‘planning’ seems to be affected by dopamine levels though it is not fully understood how. Memory plays an important part and if we monitor the way people look at a face, we see that subjects concentrate on specific meaningful areas. The superior colliculus seems to control the way the eye circulates around these. But outward behaviour tells us little about the ‘intelligence’ of a system. We might, for example, be in an airplane in which there is both a human pilot and and autopilot. On normal flights the autopilot does everything our skilled human pilot does but which one would we trust in an emergency? Suppose we have a stack of four different coloured blocks and are asked to build them in a particular different order. The search space (or the total number of different arrangements) for a computer analysing all the possibilities is 85, yet a person can envisage a solution without considering all these. In fact a test for the early onset of Parkinson’s is to give the patient this sort of problem.

The Crick and Kosh hypothesis is that the early visual system deconstructs the world into ‘features’ which are reconstructed by the prefrontal areas of the brain. Zeki and Bartels postulate that physically separate and functionally specialised sites of the visual system contribute independently to conscious visual sensation in a way that does not require neurally mediated binding or synchronisation between these areas. Igor pointed out that this does not take account of the part that memory might play in determining eye scanning. The primary visual cortex, the superior colliculus and various motor mechanisms are physical entities that contribute to ‘depiction’ in our minds but memory seems vital for coherence. There are ‘gaze-lock’ cells and cells that respond only if the eye is pointing in a particular direction. Such observations help us to understand how a moving eye can give us the perception of a solid world. But the Zeki and Bartels hypothesis doesn’t tell us how we can visualise a face or an object when we’re looking at a blank sheet of paper. We can imagine an image as being ‘out there’, not in our heads; a visual memory which may not be as accurate as an actual perception but it can none the less be conjured. Defining the ‘intelligence’ of a system is therefore tricky, it cannot be done in terms of behaviour but only somehow by the processes that go on inside it.

Training Robots

Building robots that ‘see’ the world and then respond to it in certain ways can be simply achieved by a putting a ‘look up’ computer program inside their heads. Given a certain kind of specified input the computer has to look up an appropriate kind of output. The computer can be programmed to classify experiential inputs within certain prescribed limits and to give set responses. The more the robot is trained and the richer its program the quicker and more discerning it gets.

A single neuron functions a bit that way. According to a flow of programming information it responds to recognised patterns and does its best with things it hasn’t seen before. Stuart Kauffman’s early model of a complex system had neuron-like nodes but they only had two inputs. When he put a lot of these nodes together and interconnected them everybody thought that the system would simply go for a random walk through its space of states (all its possibilities). But, depending on the interconnectivity, he found it exhibited much more stability than anyone had thought and he wrote a paper in 1969 on this unexplained stability. If we put together nodules which have stable properties then some ways of interconnecting give no behaviour at all but others give all kinds of behaviour. Complexity theory examines this kind of phenomenon.

If we consider the ‘stacking’ problem considered earlier, a single neural node is trained by receiving information about the way people restack objects. It cannot solve the problem in the way we solve the problem but it has a way of representing it or part of it which is very different from the ways people have thought in classical computer science. A representational neural network of about 180 by 180 neurons has been used to investigate this process. Each neuron responds to an input by giving out an electrical signal displayed on a visual display unit as a single dot. The system has a built-in time sequence so that changes are shown as a series of frames. The system is trained by changing a visual target to which it is responding, much like taking a series of shots of someone moving the blocks. It is made to respond to the sequence of events over and over again, reinforcing the input to each individual neuron node. If neurons are confused about their input they respond by giving odd flashes. If the characteristic of each neuron is changed by reducing its power of generalisation or classification it becomes less tolerant to changes in its input and the system behaves rather like a patient who has been given an anaesthetic. The nice sequence of backwards and forwards steps in the block building disappears, the system exhibits more and more ‘noise’ (random behaviour) and then goes to sleep. This is typical of a complex system in which the local activity of a large interconnected mass of things is changed resulting in a dramatic difference in global behaviour.

Computers that can only remember things that they have ‘seen’ like a yellow triangle or banana are considered ‘unevolved’. Human beings are able to imagine a blue banana even if they have never seen one. Some process in our brains must enable us to do it but nobody is quite sure what. Zeki and Bartels tell that visual consciousness may give us the impression of one thing though lots of different parts of the brain are active. The superior colliculus moves the eyes around though memory somehow constrains the movement. But why would a particular image or part of an image be reinforced in memory and constitute some form of ‘attractor’ in the first place? Some process of generalisation or classification seems to be at work and in mechanistic terms we can think higher levels of representation as not being all that different from lower level ones. If there are ‘natural kinds’ in the world and not just in our heads then ‘classes’ are something we discover. For example, apples, bananas and plums fall into a general class ‘fruit’. Such an increase in logical space and the exploration of it would offer a route to blue bananas. After all if plums are blue and bananas are yellow and both are members of the class fruit then the step to blue bananas seems quite small. We can represent a state space activity in which the attractors are bananas and a larger one in which the attractor is ‘fruit’. But such ideas of classification are probably far too simplistic for understanding how we can conjure up imaginary objects.

Complex Thinking

What takes us from a simple assemblage of neurons to something as massively powerful as the human mind? Igor has models in which there are about 120 different varieties of neurons though the programming of each is just a few lines. The interesting thing is that even when individual neurons that only perform fixed functions are put together very complex emergent behaviour results. By putting in information about the way the neurons might work in 'seeing' an object a simulated superior colliculus becomes sensitive to 'high spatial frequencies' and creates the vectors for eye movement. But though in very simple terms the neuron basically just looks at the similarity between things its been trained on and novel images there is much more in the process of learning about an environment and especially the part that language might play. Real neurons are funny things, when looked closely they are enormously complex; veritable chemical factories responding to and sending out signals. Some produce electrical impulses but others squirt chemicals around the local vicinity. They're very complex in the manner in which they work but it is possible to put a 'box' round them and describe in mathematical terms what they do and then it becomes simple. The way they do things is complicated but the result or function is simple.

Which is not to say what comes out of the inconnectivity is simple. The notion of an 'attractor' is a mathematical one, coming from 'trajectories' and the like. Point attractors are where trajectories home in on some 'space'. But though recurrence in complex systems can be impossible to express in mathematical terms there do seem to be important attractors and feedback in the behaviour of organic systems. 'Feedback' is a term which should not perhaps be used in the context of complex systems because it tends to elicit the question: 'is it positive or negative?' which comes from 'control theory' which treats stability and instability very simply. You cannot apply the term to a system of neurons or nodes in which some are acting in an inhibitory way and others in an excitatory way. Anyway in order for a system to have well behaved attractors it must be complex.

How low level activity in the brain might lead to higher level abstractions is the subject of much debate. If we are trying to describe a system as outside observers then we might distinguish low level activities and higher levels of abstraction and some kind of feedback acting in a corrective manner. But what 'getting it wrong' in terms of system change is hard to say. Putting neurons in a new environment makes things go very 'noisy' and unstable

'Changing ones mind' can involve substituting one image for another and this suggests a non linear process. Neuro-anatomy also indicates neuro-chemical activity has a lot to do with mood or imagination and people whose brains have localised lesions in areas which are implicated in emotional states seem to have greater difficulty in making decisions than those in areas associated with logic or geography. And this approach may be important. A robot on Mars might well benefit from some equivalent of fear or pleasure to achieve its mission and ensure its survival. But changing ones mind means more than simply substituting one image for another. It may also involve taking a personal decision and this would involve the notion of 'self'. It was suggested that Ross Ashby's idea of homeostasis or maintaining a stable position in a changing environment might be of help. Some kind of internal training that allows regeneration of the idea of self would allow measurement of pain or pleasure. Igor's stance, however is to see anxiety more as something to do with restructuring the state spaces that are in place. A robot would say to itself: 'hey I'm not making the right predictions here!' He (or she) would not feel right so something would change. Since the 'noise' level is an indication of confidence in the system it would go up. When building robots it may be important that it can run away when it doesn't feel right but there's a long way to go on that.

Changing Metaphors

Inman pointed out that in 1951 the most powerful computer in the world cost about four million dollars now the equivalent power costs about forty cents. Moore's law predicts that the numbers of transistors on a chip doubles about every eighteen months and this exponential growth has meant that the computer in a watch is as powerful as the one built in 1951. But we do not exploit this power that we have and this is as true for practical engineering as it is for the designing of robots.The reason that we fail to fully exploit the power that we have is that we evolved to swing from trees and fun after things on the African plain rather than design complex systems. We really aren't very good designers or mathematicians in that respect. We build heuristics which enable us to understand the world by carving it up into modules that we make as independent from one another as possible. This has been the classical approach to building computer. A computer is very good at solving our mathematical problems and because they're good at that we've tended to use it as a metaphor for the brain. This is no different from previous times when the steam engine or the telephone network was used as a metaphor because they were the most complex things at the time.

'Artificial intelligence', now better known as 'computational intelligence' has had some enormous successes. The ability to manipulate text on the computer screen was once regarded as AI but since it is now so commonplace nobody would regard it as such. But the approach falls down in the pursuit of practical robotics because it is based on the assumption that all cognition is about finding out facts about the world, putting them into a computer and manipulating them like symbols in a mathematical equation. Classical computing can do calculations in a flash that would take the human brain an hundred years but it's no good inside a robot kitten that has a sporting chance of crossing the road without getting run down. A classical robot called 'shaky was once built to observe where some solid cubes and pyramids were in a room, build up a picture in its head and then plan its move. It took 15 minutes to do that and move a few centimeters. If an object in the room was moved it had to start all over again. This was not due to limited power in the computer but a 'mindset' that constrained the kind of architecture that the designers put in the robot's head.

Such architecture is known for short as 'sense, model, plan, action'(SMPA). Inside the robot's head there is a module which takes sensory input such as light falling on light cells and builds up a model which can be put into programmatic form.The question 'how do I get around the world?' which has been converted to a bit of logic then becomes answered by a few algorithms and the solution is transmitted to the motors as action. There's the whole Church-Turing assumption that any process that a human being carries out is like a list of instructions. The moon takes a certain pathway through space but it is people that compute its trajectory. The moon simply does whatever it does though how we might know that is a tricky philosophical question. Assuming that the brain operates the same logical rules that we use in our language is a bit like assuming that our calculation of its trajectory is the one that God made. Of course we can use a computer to simulate what the moon does and we can use a computer to simulate a nervous system or even evolve robots whose brains are evolving dynamic systems but we don't have to believe that reasoning is the prime consideration in building a walking robot. Inman's point is that the math in building some sensory coordination, about how we manage to put one foot in front of the other without falling over is pretty tricky but natural behaviour such as walking can evolve in a body with no brain. A model of a leg which can be simulated on the computer or made in real wood, providing it has a constrained knee joint and a universal hip joint will walk down a tilted plank in a very human like way. Inman calls this the 'dynamic systems' approach to robot building. Rather than assume reasoning 'all the way down' Inman would like to assume dynamics all the way up. His suggestion is that not only is this approach appropriate to low level motor sensory mechanisms but that even our ability to calculate two times three is equal to six is just a more sophisticated example of dynamic skills. When dynamic systems in the form of nervous systems are coupled with the dynamics of the environment, appropriate behaviour for survival results. We can describe simple systems in terms of variables in an equation and we think that any physical system we set up with pendulums, cogs and springs is in principle governed by physical laws that apart from inevitable 'noise' we discover. Interaction of an organism with it physical environment is via sensors and motors and what we rapidly learn is that if we understand one dynamic system on its own and couple it with another one we understand, the resultant system is full of surprises.

We and cats are designed by evolution and AI design has so far not taken this into account when building robots to deal with the entirely new. Whereas a bacterium is truly adaptive and can hunt down nutrients and avoid harmful chemicals our classically designed robot cannot. This largely stems from our belief that the way to deal with the world is to build some kind of working model of it in our heads. And this in turn arises from our stance as independent observers. Such dualism separates the mental from the physical. It assumes there is a way the world is, in an objective sense and the scientist or robot is a spectator observing it through various sense organs. Though we no longer regard the brain as storing little pictures of the world, neural network designers still seem to assume that building a brain means somehow making the connections that produce the attractors that model the 'outside world'. Just as it is deemed necessary to have certain specific genes to have certain characteristics it becomes necessary to have certain attractors in the neural network system. Inman seems to think differently. He denies that he has any models in his head whatsoever. He admits that we all use symbols but that they are not represented in our brains as any particular patterns at all. The whole ideal of the brain as a computer implemented in 'wet stuff', he feels, reaches back to the television screen in our heads or the little 'homunculus peering out at the world.

A 'Dynamic Systems' Way To Develop Robots

Trivially we have brain states but the patterns in our brains do not tell us anything about how we think. Darwinian evolution requires heredity, variation and selection and DNA holds coded instructions in a manner that ensures the replication of an organism. If the instructions should be completely garbled the organism would not be procreated.

Suppose we try to evolve a paper dart by writing down some folding instructions in a completely random manner. We could produce a number of variations, throw the products out of the window and see which went the farthest. We then look for some pattern or meaning in the sequences rewrite the list incorporating the pattern and repeat the experiment. By doing it enough times we stand a good chance of producing a reasonable paper dart.

As robot engineers we want to produce robots that are adaptable to environments we specify. We don't much care how their brains evolve but we do want them to exhibit some particular behaviour patterns. This is unashamed behaviourism. By making a simulated neural net with sufficient connectivity and nodal instructions to enable 'loops' and 'attractors' to arise, we simply evolve the system in the way we want. Since the robot is a system coupled with a physical environment we also have to build in the ability to alter form and locomotion. Of course this isn't like natural evolution which involves the co-evolution of an organism in a whole ecosystem though specific trends such as 'arms races' between species can often be identified. Nothing tells lions, for example to run as fast as they can to catch antelopes or antelopes to run as fast as they can to escape being eaten but the end result is that we end up with fast lions and antelopes.

Using some kind of artificial DNA that encodes a dynamic system is not simply computing outputs from inputs. The unknown evolved architecture may involve all sorts of attractors and enhancing or inhibiting nodes but it remains like a brain in a vat with its own internal unknown dynamics. It's really just a convenient way of using evolution to pick a way through logical spaces by dealing with different environmental pressures. A typical setup involves an evolved net connected to sensors on one side and locomotors on the other. The robot is then put in an environment which is real or simulated, put through the test and its score card marked. An eight legged robot was produced in this way with locomotors and infra-red sensors that with more sensitivity would potentially enable it to cross the road without being run over. Its brain was completely produced by artificial evolution and downloaded after some ten thousand generations. But it had nothing inside its head that we would wish to point to as a model of an ‘external’ world. Its brain was simply a ‘black box’ which we could use to generate the behaviour we required. Biologists define ‘adaptability’ in an a posteriori way whereas robot engineers can aim towards particular goals. But it’s still a bit like gardening. Mutual adaptation does not build ‘Frankenstein’ monsters but it does build robots that want to live and given a choice between an autopilot and an educated robot we might be wise in certain circumstances to choose the latter.

Conclusion

We can approach the subject of artificial intelligence in two ways: as a means of understanding what might go wrong in our brains either physically or psychologically or as a means of producing intelligent robots. The term ‘Dynamics’ takes us towards behaviourism whereas the term ‘cognition’ leads us away from behaviourism to contemplation. Human learning has much to do with contemplation. It is working out whether action A will be right or some other action B will be appropriate and a great deal of representational experience is involved in this. Our ‘fitness’ depends on our ability to detect things not in the real world but as what might be called ‘opinary emanations’. As behaviourists we might see such imaginings as merely rehearsing procedures that have evolved throughout our history though how that ability could have resulted purely from the ‘dynamics’ is hard to see. The ability to depict a blue banana, for example, seems to point to something else.