Constructivism in AI: Prospects, Progress and Challenges

Frank Guerin[1]

Abstract. This position paper argues the case for the application of constructivist theories to Artificial Intelligence, with particular emphasis on Piaget's theory. The idea of building an artificial baby is an old one in Artificial Intelligence, yet it is difficult to execute because little is known about the information processing mechanisms which babies use to learn. That which is known comes from non computing disciplines and has not been exploited very much in Artificial Intelligence. Part of the difficulty is that many AI researchers do not know enough about these other disciplines; another difficulty is that where AI researchers do know some of the theories from other disciplines, many do not see their value. This paper tries to make a case for the value of Piaget's theory in particular.

1 INTRODUCTION

This paper sets forth a research agenda for the application of constructivism to Artificial Intelligence. We take the view that there is a mine of useful ideas in constructivist theories, particularly Piaget's[2], that have yet to be fully explored by AI researchers.

Section 2 gives a brief description of what is meant by a "constructivist approach to AI", and contrasts it with non-constructivist approaches. It uses the behaviour of "groping with a stick to retrieve an object" as an example task, to contrast the two approaches. It then gives a fairly detailed account of how this behaviour could be acquired by an AI system which follows Piaget's theory. This account serves to give the reader a detailed picture of what is entailed by the constructivist approach. Section 3 makes some arguments to justify why we believe the constructivist approach to be worthy of investigation, and tackles some common objections. Section 4 reviews some of the existing work in constructivism in AI, to see what has been achieved and what remains to be done. Section 5 briefly outlines some of the major challenges to be tackled in following the constructivist research programme advocated here. Section 6 concludes.

1.1 Controversies about Piaget's Theory

There is considerable controversy over the claims Piaget has made, particularly when it comes to what knowledge is innate or learned. Some results in the 1990s purported to show a great deal of innate knowledge, far beyond what Piaget had proposed. However, researchers in the constructivist camp followed up with studies of their own and drew different conclusions. The controversy is a lively one, with claims of improper experimental procedures, overinterpretation of results, failure to replicate reported results, etc. This paper will not delve into the controversy, interested readers could consult Cohen and Cashon [5] and Haith [8] as a starting point. This controversy looks unlikely to be settled in the near future; in the meantime, we must concede that it is not certain that constructivists' theories of human learning are an accurate account of what humans actually do, but we take the position that it is worth trying it out for AI. Apart from possibly helping us to build more intelligent programs, an AI investigation might also shed light on whether or not constructivism is a viable theory to explain human learning.

2 CONSTRUCTIVISM IN AI

This section explains what we mean by constructivism in AI. We use an example behaviour from Piaget's observations, and explain how a constructivist approach to AI might go about implementing it, mostly following Piaget's theory of constructivism.

Piaget describes the "behaviour of the stick": the infant seeks to take possession of an object which is located out of arm's reach; the infant uses a stick as a tool to draw the object into the range of his arms, and then takes possession of it. To keep things simple we can restrict our attention to the case where the infant is provided with the stick by an adult. In this case the infant typically becomes capable of the behaviour of the stick roughly between 12 and 18 months (Piaget's fifth sensorimotor substage). The case where the infant has not been given the stick, and must think of the idea to use it himself, is more advanced and belongs to the subsequent substage. Note that being provided with the stick, and being shown an example of the required behaviour, does not at all make the task trivial for the infant. An infant who has not yet achieved the behaviour of the stick may attempt to copy the adult behaviour, but can only do so coarsely and will merely strike the desired object, and be unable to draw it closer; the infant at this stage will limit himself to repeatedly striking the object in the same way, and will not attempt to modify the way that the object is being struck. This primitive behaviour may persist for several weeks before the infant properly begins to grope and thereby acquires the behaviour of the stick. By groping we mean that the infant modifies the way of striking, and so identifies the type of striking that brings it closer.

We will now contrast two possible approaches to implementing the behaviour of the stick in an AI system: firstly a non-constructivist approach which makes use of prior knowledge, and secondly a constructivist approach where the infant must construct the relevant knowledge. For the non-constructivist approach we could employ a reinforcement learning algorithm. The artificial infant should get a small reward whenever the stick is brought closer to the object; this could allow the infant to quickly learn to bring the stick in contact with the object. Upon contacting the object with the stick, the infant should get a large positive reward whenever the object is moved closer to the infant, and a negative reward whenever the object is pushed further away. Through random groping the infant will find the actions which draw the object closer. This is all doable with present

[1] University of Aberdeen, Scotland, e-mail: f.guerin@abdn.ac.uk

[2] This paper draws heavily on Piaget's account of infancy, as described in his trilogy of books [10, 12, 11]. Where specific references are not given, we are referring to this trilogy, and usually to the first two books [10, 12].

techniques, and will lead to the artificial infant retrieving objects with reasonable efficiency. The main "cheat" being exploited here is that we have told the infant (via rewards) that it needs to bring the stick in contact with the object, and that it needs to take certain actions which make the object come progressively closer; we have also provided the infant with information about when the object is coming closer, so it does not need to worry about knowing how far away the object is. All of this amounts to giving the artificial infant prior knowledge about the world, and specific knowledge about how to achieve a particular goal.

The constructivist approach would try to avoid giving the infant any prior knowledge beyond that which is absolutely necessary to bootstrap the learning process and allow the infant to learn in reasonable time[3]. Apart from this minimal innate knowledge, the constructivist approach aims to allow the infant to create the required knowledge for itself. For example, innate knowledge might include the ability to grab that which touches the hand, the ability to suck that which touches the mouth, the ability to make random arm movements, etc. By bootstrapping from these initial abilities, the infant must learn how to suck the thumb, how to grab and suck world objects, and how to interact with world objects in more complex ways. Through this interaction the infant must somehow learn higher level knowledge about the world, gaining knowledge of space and objects and how to manipulate them. This would eventually lead to the behaviour of the stick.

The path along which an infant develops the required knowledge has not yet been determined by psychology (i.e. what specific episodes lead to the acquisition of what specific knowledge). Piaget's theory gives a sketchy overview, but many details remain to be fleshed out. Coming up with *(1) a plausible path of development* is the first challenge for a constructivist approach to implementing the behaviour of the stick in an AI system. A speculative and somewhat sketchy account of such a path is given in the next subsection. Given such a path the next step is to *(2) computationally model the information processing*, i.e. code learning algorithms which can profit from the experience specified by episodes in the path, in order to learn the required knowledge. How this can be programmed is not at all clear with present day AI techniques. This is the second challenge for constructivism in AI. Section 4 reviews some related work which has developed algorithms for some types of constructivist learning.

2.1 A Possible Development Path

This section gives a highly speculative account of how an infant (real or artificial) might develop the behaviour of the stick, starting out with only some very basic innate abilities. It follows Piaget's theory very closely, with some minor variations to explain (roughly) how a computer system might learn it. It is speculative because it has not been tested either with real infants or AI simulations. This is the case with Piaget's books on infancy; he proposed a theory which accounted for what he observed in his own three infants, but he did not attempt to falsify it through larger studies. It must be conceded that the description is also sketchy in places, and many details would need to be worked out to create a computer implementation. The purpose of presenting this fairly detailed account here is to give the reader a clearer idea of what we mean by a constructivist approach in AI, and to highlight some of the important aspects of the constructivist point of view. Whether this account is correct or not, some description along similar lines could be proposed and tried out.

The acquisition of thumbsucking seems not difficult to account for: when the infant notes that his hand has come into contact with his mouth, he can remember the movement which brought it there from some near location x, and so repeat it in future, if he finds his arm at x (he can recognise this location x by proprioception). In a similar way he can find the action which brings his hand from some more distant location to x. Extending this idea through all the space in which the arm moves, the infant can find actions which bring his hand to his mouth from any initial arm location. With this ability acquired, one can see how he can learn to grab a touched object and take it to his mouth. One can also imagine how this can be extended so that he learns to focus on some object with his eyes, and then to move his hand to the centre of his vision to grab it, and then suck it. This is a significant development; it leads the infant to observe new interesting phenomena; new unexpected things will happen, because he is now interacting with the external world (as opposed to simply sucking a part of his own body). For example, when reaching with his hand to grab a hanging object, he may accidentally knock against it, producing a swinging motion. This interests the infant and he tries to rediscover the action that caused it. In this way he discovers a motion which can effectively swing the hanging item; he refines this movement to develop an effective "striking" action, to augment the swinging of the item. This strike action, being new, is now tried out on every object the infant encounters. The infant is generally eager to try our any new actions as much as possible. A number of new actions are developed at this stage, but we can restrict our attention to striking.

The effort to grab objects in the external world also inevitably leads the infant to confront problematic situations; for example, situations where the desired *object* is visible, but some *obstacle* blocks the infant's hand from grabbing it. The solution to this problem is to simply push aside the obstacle, however this is not at all obvious to the infant at this stage. There are scenarios where the infant could use his "striking" action to knock aside the obstacle, yet he does not. This reveals a lot about the infant's world model at this stage. Each action in the infant's repertoire is isolated; there is no coordination between them, and hence no knowledge of their relationships. Thus the infant's world model consists of a series of fragmentary spaces: one to describe the effect of each individual action. He knows well that striking the obstacle will move it in a certain direction. Why is it that he cannot strike and knock aside the obstacle in order to clear the path and reach for the desired object? It is because if he thinks of striking aside the obstacle, then he is considering this striking action along with its fragment of space (by this we mean that he has knowledge of the postcondition of the action: that the struck object will move away from him in space); the obstacle is represented in this fragment, but the desired object is not. So although he could foresee the obstacle moving, that "mental image" does not include the object, so he cannot foresee that the object would become accessible. To foresee this he would need to represent both the object and obstacle in his mental image. That presumes a more sophisticated representation of space which has connected up the two fragments. In fact this "connecting up" of space will happen through the coordination we are about to describe. By chance, on one of the failed attempts to grab, and upon clashing with the obstacle, the infant may decide to strike the obstacle, without foreknowledge of what this will achieve.[4] Having done this, the infant has not forgotten his recent goal to grab the object,

[3] We do not want to make the learning problem so difficult that the time required is inordinate; e.g. it would take too long if we had a reinforcement learner that only got reward when the object was retrieved.

[4] This departs from Piaget's account which describes it as an intentional act to "negate" the inclusion of the offending obstacle in the grabbing scenario.

and now sees that the path is clear, and so grabs it. This coordination is remembered, and used thereafter. The infant also learns to discriminate the visual appearance of a scene where an obstacle to grabbing is present from one where an obstacle is absent; using this knowledge the infant is able to employ the intermediate striking action in those cases where it is required, without having to first bump against the obstacle on a failed grab. This striking action, to remove an obstacle, gets used repeatedly, and becomes refined, gradually leading to a displacing action which can pull an offending obstacle aside. With these acquisitions the infant is now able to pay attention to the relationship among any two objects which are seen, and especially to notice when one is "in front" of the other, or could be moved "in front" of the other. The acquisition of this displacing action means that he has an action in his repertoire which can change the relationship among two objects. This action is of a higher order than those previously in his repertoire; in Piaget's language: it belongs to the next stage. Having acquired this behaviour, it can again be extended: the infant can learn to bring one object closer to another, when that causes an interesting interaction between them which he would like to repeat.

Let us now look at the behaviour of the stick and see how far we are from achieving it. By now the infant could strike an object with the stick if given the stick and shown an example of the behaviour; however, he could not draw the object towards him. His limitation at this stage is that when he strikes the object, and it moves a little, he is unable to interpret this movement (the movement is meaningless to him and he is effectively blind to it). A more advanced infant would understand that if it moves a little because of the stick, then the stick has the power to displace it in space, and further groping might well find actions that move it on a trajectory towards the infant. In order to be able to use the stick as a tool to bring a desired object closer, the infant will need to know the trajectory on which the desired object should travel in order to come within range of his arms. This presumes a knowledge of where the object is currently, and the location to which the infant would like to bring the object. Surprisingly, the infant at this stage (8-12 months roughly) does not actually know where an out of reach object is, even if it is seen. Knowledge of location in space will need to be constructed. Understanding this is crucial to understanding the constructivist's hypothesis about the infant's view of the world (or knowledge of space in this instance), and its development.

Piaget describes the child's increasing knowledge of space as akin to an expanding sphere. By investigating the space accessible to his arms he comes to have a practical knowledge of this "near space"; practical in that he knows how to grab something he sees in this space, but he does not understand how different locations in this space are linked. At this stage looking at objects in far space (out of reach) is like looking at images painted on the inner surface of a sphere, which are mostly static, but sometimes move in mysterious ways. An object which moves from near space to far space does not only change its position for the young infant, to the infant it is as though it changes its nature. The infant at this stage only reaches for objects in near space and does not bother with those out of reach. With development he begins to see these objects as constant and begins attempting to reach for objects which are just out of range. When he begins to pay attention to relationships among two objects in near space (a development which could be reached by the path described above), he then begins to notice the parallax caused by motion of his head. Knowing the relative positions of two objects A and B in near space, and the parallax caused by head movements, he can make the conjecture that there is a correlation between parallax and relative position. When he then sees the parallax between a near object C and one which is just out of reach D, he can reason by analogy that the far object D is probably behind the near object (i.e. because the parallax observed between C and D is the the same as that between A and B therefore the spatial relationship is probably also the same). Thus he begins to locate the far object D in space. This development can then be extended to notice that a further object must be located behind D, and so on. In this way he can begin to locate the position of all the objects in a room relative to one another. However limitations will still exist for very far objects. The expansion of the child's sphere happens over years, with young children wondering why the moon follows them, before reasoning that it is merely the parallax motion between it and the trees.

It may be difficult to imagine the infant's primitive view of the world, given the things we take for granted from our perceptual experience. It is perhaps easier to understand the development described above by looking at how adults can also extend their knowledge of space by expanding the sphere in which distances can be reckoned. Primitive societies tend to view the stars as points on the inside of a hemisphere, sometimes the sun and moon are also thought to be the same distance from the Earth. Aristarchus of Samos was able to reckon that the Sun was much further away than the moon by considering the angle between the sun and moon when the moon is half-lit (imagine that the sun is a lantern, the earth an apple and the moon an orange, and then consider the angle made by lantern-apple-orange, when the orange is half lit from the apple's perspective) . Thus he was reasoning by analogy with his knowledge of illuminated spherical objects in the near space where distances were known to him. He hypothesised that the same laws held for far objects. In more recent times telescopes have allowed parallax to be used to estimate the distances to stars, by measuring angles from different orbital positions of the earth. This is an example of constructivist learning; it the application of known facts (diameter of Earth's orbit, geometry) together with new observations (angles subtended by stars in Winter and Summer) to construct new knowledge (distances to near stars).

Returning to the infant and the stick, we can now see how the infant could come to know the location of an object which is out of reach. He can reason relative to successive landmarks, starting from some location in near space, and going to the desired object, each being behind the other. When the desired object is struck, the infant can note any movement it makes relative to nearby landmarks (even elements of a pattern/texture on the floor could serve as landmarks). This relative movement is now salient; in contrast to the infant at the earlier stage, any relative movement between two objects is now interesting, because the infant knows an action which can do that, and has played with changing relative positions. The infant can understand that if the stick can cause a small relative movement, then a movement can probably be found to bring it closer relative to some nearby landmark; this could then be repeated to bring the object in front of successive landmarks and thus into near space.

An objection may be raised here to say that the infant does not need landmarks to precisely locate objects in space, because he could gauge the distance to which his eyes focus and the direction in which they point. This would require that the infant had carefully calibrated his eyefocus to match it against different distances, and had noticed the analogy between changing focus and changing parallax (for example), to conjecture a relationship between changing focus and distance. This is a possibility, though it seems unlikely; it would be difficult even for an adult to reliably tell if an object has moved closer or further away when it makes a small movement in some random direction in a featureless space (contrast with the object moving on

a patterned carpet). Similar arguments and counterarguments can be made for attempting to locate an object in space based on its apparent size.

This then is what we mean by a constructivist approach to learning the behaviour of the stick. In particular constructing the knowledge of space requires quite a sophisticated type of reasoning by analogy. Constructivism proposes that the kind of learning machinery which can do such analogical reasoning is innate in the infant. Thus a constructivist approach to AI escapes the burden of coding a great deal of innate knowledge, but takes on the burden of coding a very sophisticated learning algorithm. The scientific objective of constructivism in AI is to discover what type of information processing mechanism could implement the constructivist theories of Piaget and others.

3 PROSPECTS: The Case for Constructivism

Having introduced the basic ideas of what we mean by a constructivist approach in AI, we now make a case for why it may be a promising approach. We will support the position by considering some arguments for and against. In comparing the constructivist and non-constructivist approaches to learning the stick above, one may be struck by the extremely poor level of initial knowledge of the constructivist system, lacking even the ability to tell where visible objects are located in space. It is reasonable to wonder if there is anything to be gained by handicapping the system so severely. The counterargument: the point of the proposed approach above is really to advance our scientific knowledge of constructivist learning; our ultimate goal is not to achieve the behaviour of the stick, else we would code the required competence directly. By handicapping the system in terms of prior knowledge in this task we are forced to come up with a system which can gather the required knowledge through its interactions with the environment. Thus we are hoping that by making the task specification of the problem roughly similar to that which faces the infant, we may be able to come up with an algorithm which has some similarity (at a high level perhaps) to what the infant uses to solve the problem; hence it should be extensible and go beyond the specific tasks which it was trained on.

Our interest in constructivist learning is not purely to advance our knowledge of cognitive science: we expect that constructivist learning mechanisms will prove useful in practical AI systems; this position will now be supported. We expect that endowing AI systems with constructivist learning mechanisms will bring two benefits: (1) it is likely to be a good approach to the commonsense knowledge problem (i.e. let the program learn itself with constructivist mechanisms); (2) it is likely to be a good approach to allowing systems to generalise from what they know and so learn how to cope with new situations (in fact it treats these issues as central).

Point (1) claims that it will be easier to build a constructivist learner and let it acquire the commonsense knowledge of a three year-old (for example) through interactions with the world, than it will be to code that knowledge directly. A counterargument could be that perhaps constructivist learning is just some idiosyncratic way in which the child seems to learn. Given that we know a lot more about objects and space, perhaps we could code our advanced knowledge directly into our programs, and then they would have no need to perform constructivist learning. This counterargument has some force, because if we consider constrained tasks, then constructivist approaches are certainly not the best. For example, industrial robots can perform highly skilled operations which require some knowledge of space and objects, and these robots do not have to go through a long apprenticeship, as the child does. Industrial systems often achieve tasks which humans could also do, but in a very different way. Their representation of space, for example, is clearly very different to that which would be built by constructivism, yet it works well for the tasks these industrial systems have to do. However, experience has shown that, for general knowledge, it does seem to be exceedingly difficult to code all the commonsense knowledge of even a child into a computer. Existing approaches which have been successful in constrained domains have not proved to be extensible to general knowledge. It is worth trying alternative approaches.

It is not clear if constructivism is the best approach to learning about the world (and space for example), but it is one way that works (in the human). One of its promising aspects is that it does seem to give a plausible account for some of the human abilities which current AI systems lack, and some of the idiosyncrasies of human reasoning. Humans often make incorrect analogies, as can be seen from the history of science and pre-scientific notions. Consider the following gem from Piaget's investigations; it is a child's response to a question about why a helium balloon goes up: "Because there's a gas inside, when there's a lot of gas it's heavy, it's very strong and then it flies." Though the constructivist mechanism can sometimes lead to wrong conclusions, it is perhaps more surprising how often it leads to useful conclusions. The child quoted clearly has some intuitive notion of inner force or strength, and though this is for the moment confused with weight, it is nevertheless a concept which will lead the child to correct deductions in many cases. It is precisely this type of intuitive concept which present AI systems are sorely lacking. The constructivist approach we advocate forces us from the outset to find representations for knowledge which are extensible, and which facilitate analogical reasoning (otherwise the development path we attempt to model will be unachievable), thus the hope is that we will come up with representations for intuitive concepts such as this. In contrast, in a non-constructivist approach we would focus on achieving a specific competence, rather than modelling a development path. Then we code in knowledge directly into the system, and the danger is that we do it wrongly; i.e. we represent things in a way which is good for that particular competence, but not very useful for performing general tasks, for drawing conclusions or for extensibility.

As for the usefulness of the knowledge coded, it is interesting to note that in the Piagetian account of the construction of space, it is constructed by assembling fragmentary spaces, where each is a fragment of knowledge about a known action. The upshot of this is that when the space is constructed, and when the infant sees a distant object, he immediately knows the actions which could manipulate it relative to some nearby landmarks. (He may not yet know the behaviour of the stick, but he knows that if he were within reach of the object he could manipulate its position relative to those landmarks.) He will know an action to move it closer relative to some nearby landmark (that is the action of the hand drawing an object closer relative to a landmark). In fact the object's position is effectively represented in terms of a series of actions which could bring it to the infant, or alternatively, the moves the infant needs to perform to get there. Piaget cites a pertinent quote from Poincaré to support his theory about the conception of space: "to localise an object merely means to imagine the movements which must be made in order to reach it". More generally, to perceive a scene, according to Piaget's theory, is to simultaneously be aware of all the different actions that could be applied there. This is because the construction of the actual objects perceived is performed by recruiting a host of low level fragments of sensorimotor knowledge (i.e. construction of a perceived object from the image sensed). In looking at a distant building, all that is sensed are a few points of light on the retina, but the mind elaborates this so

that a three dimensional building with inside, outside, rear, etc. is perceived. This elaborate 3-D structure is constructed in the mind by making use of our past experiences when we walked inside buildings, upstairs, around the back, etc. and indeed many more general experiences which have discerned the solidity of materials, dimensions, positions in space, etc. [10, see especially p. 189-190]. This is what Piaget means by the construction of reality. It explains how human visual perceptual experiences could be quite different from standard approaches to computer vision (see also [14]). Computer vision seeks to reconstruct the 3-D structure that is being viewed by 2-D cameras, and to recognise objects there. However the human version considers all the past actions that were applied on similar shaped surfaces, and it is largely in terms of these actions that the 3-D representation of the scene is reconstructed. Therefore, when the human seeks to manipulate what is seen, to achieve some goal, the objects and actions that could achieve that can immediately spring to the forefront for consideration. In this way the constructivist approach to acquiring world knowledge holds the promise of answering Sloman's concerns about "affordances", and the *understanding* of surface structure [14].

The idea of advanced knowledge being built on top of simpler well tested knowledge (i.e. sensorimotor schemas at the lowest level) applies to all constructivist acquisitions. The contrast between a constructivist knowledge representation in AI, and a classical AI representation to solve the same problem is similar to the contrast between a student rote learning how to perform a particular mathematical operation, or a student understanding the operation as a new combination of operations he has already mastered. In the case where the student can construct the new operation as a function of known simpler operations, the knowledge is much more useful, and can be adjusted and applied in diverse situations; in contrast, the rote learnt procedure would be rigid and only useful in a narrow set of situations. Piaget's statement that "to understand is to invent" is pertinent here. To take this argument back to AI, one could criticise attempts to code knowledge such as naive physics into an AI system, because a system with such knowledge would not understand how to apply it, when it is not grounded in its own more primitive sensorimotor knowledge.

The argument in support of point (2) above continues the idea of analogy and extensibility introduced to support point (1). In the case of (2), as opposed to the commonsense knowledge case, we are considering scenarios where the knowledge required may be unknown to the system builder at design time, so having the designer coding the knowledge is not an option. Constructivism makes no distinction between learning commonsense knowledge (1) or generalising from known facts to learn to cope in a new situation (2). They both require the learner to conjecture a new theory and to test it, and refine it as necessary. The same mechanism is used in both cases. Thus the constructivists' claim is that the machinery for generalising and analogy forming which is required to form a solution to a difficult new problem is one and the same as that required to learn the basic world knowledge which children learn. The claim is that infants are doing scientific discovery all the time. Examples from the study of scientific discoveries provide support for the constructivist claim; analogy with previous situations seems to be the key to conjecturing new models of how the world works. It would be difficult to provide any alternative account, because a human can do nothing other than conjecture based on things already known. To conclude this argument: It is impossible to say if constructivism is the best way to learn, but there do not seem to be any viable alternative accounts of how this type of knowledge acquisition could proceed.

3.1 A Common Learning Mechanism?

An objection is sometimes raised over the constructivist claim that there is a common learning mechanism which is being used by the adult to discover new knowledge, and by the infant to learn basic world knowledge. The objection is that surely the adult has a superior ability when it comes to adding knowledge; surely the functions the adult uses are different to those of the child. The adult has a logical framework in which hypotheses can be entertained and dismissed, and contradictions can be noted. The child is often comfortable with giving two somewhat contradictory explanations for a phenomenon. For example: large boats float because they are very heavy and strong and can push down the water, while stones sink because they are heavy and strong and can push apart the water to get to the bottom. The response to this objection is that the adult has more learning tricks in addition to the core learning mechanism which is common with the child's; these extra functions of the adult were in fact built by the core learning mechanism.

The same phenomenon can be seen in evolution. The basic mechanisms of evolution are very simple: replication, mutation and natural selection. But as evolution progressed it evolved fancier tricks. For example, there are two types of genes: structural genes and regulatory genes [13, Ch. 20]. Structural genes code for "building block" proteins for example, while regulatory genes control the expression of other genes. Regulatory genes can act like switches, controlling structures that appear on the body. The evolution of genetic switches makes subsequent evolution easier. It can be thought of like a computer program becoming modular, where modules can be called with parameters; it makes modification easier. Sexual reproduction is another example of a fancy trick by which evolution has been able to accelerate its own learning. These tricks are not present in early life, and do not need to be; the basic mechanisms of evolution are enough, the other tricks will appear through evolution. In the same way, constructivism claims that many of the fancy learning tricks used by adults do not need to be present in the innate learning mechanism.

3.2 Why Copy Infants?

An objection may be raised over the choice to copy acquisitions in infancy, as opposed to acquisitions in childhood or adulthood. Piaget's theory hypothesises that the same learning mechanism is in use at all ages; given that we are really interested in modelling the learning mechanism, rather than a particular acquisition, we could well model some learning episode in an adult. The UK Computing Research Committee's Grand Challenge 5 meeting reports that "it was argued that newborn infants are much harder to study since most of what they do is very inscrutable" [5]. It is true that infants are difficult to scrutinise because you cannot ask them what they are thinking, however children and adults could be viewed as equally inscrutable because they have subconscious processes which are not open for introspection, and sometimes they can give very inaccurate accounts from introspection. An argument can be made that infants may be more scrutable, because at a young age they are incapable of simulating a possible course of action in their heads, and must try it out by groping in the real world to see what happens. Furthermore, at this stage they may use their body to physically represent a situation they are dealing with. Piaget recounts an observation of one of his daughters who is attempting to put a chain into a matchbox which is open just to a 3mm slit. Not succeeding in making the chain enter, she seems to represent the slit with her mouth, opening her mouth

[5] http://www.cs.bham.ac.uk/research/projects/cogaff/gc/

wider and wider, and thus discovering a solution to the problem, i.e. to enlarge the opening, which she promptly does. After 18 months the infant begins to solve these types of problems covertly, simulating them inside his head. From this age the infant certainly seems to become more inscrutable, when many ideas are experimented with internally, and the infant suddenly comes up with a solution to a problem, seemingly out of nowhere.

More than the above argument though, the main argument in support of modelling acquisitions in infancy is that it should be easier, because knowledge is being built on a smaller set of existing knowledge. The acquisition of new knowledge requires finding analogies with existing knowledge, so we need to trace out exactly how each new acquisition is built on previous knowledge. This can be very complicated; especially when there are so many possible pieces of existing knowledge that a new piece of knowledge could relate to. Our task (in pursuing the constructivist approach to AI) is to come up with the mechanism which can add the new knowledge. We need a precise specification of this task in order to write an algorithm to do it. The task is: given a certain starting state of knowledge, and a new experience which goes beyond (or is not consistent with) existing knowledge, add some new knowledge which accounts for (is consistent with) that experience. One of our challenges is that it is difficult to know the starting state of knowledge, and having an accurate description of this state is crucial, because if we miss something, or add too much, we could make the task of addition too hard or too trivial. This is why we advocate going back to infancy, there the starting state of knowledge should be simpler than at any other age.

The problem of knowing the starting knowledge state does not go away entirely however, determining the innate knowledge of the infant is very difficult. Psychology experiments can be helpful here, in particular the habituation technique. This technique presents an infant with a familiar scene A until the infant is bored of it and looks away, then some other scene B is presented. If the infant continues to be bored by B and looks away, then one can conclude that the infant notices no difference; if on the other hand the infant looks longer at B, then one can conclude that the infant notices a difference. An example of where this can be applied would be showing an infant an object entering and emerging from a tunnel, and then showing the infant one object entering and a different one emerging (younger infants will notice no difference). Furthermore, to support the argument that infants are scrutable (albeit via rather elaborate experiments) an experiment carried out by Bower [1] is most interesting. He separated two groups of infants, and gave one group special training sessions, which involved having them watch scenes of objects pass inside other objects and re-emerge, for example. Simply viewing these displays accelerated the infants' cognitive development and led to improved performance on manual tasks also. The group that received the training reached the next Piagetian stage; the untrained group had not reached this stage in the same time. This experiment is an excellent example of the kind of psychological investigation which could complement a constructivist AI programme. The experiment clearly identifies the precise experiences which have led to the construction of a particular stage of more advanced world knowledge.

Finally, a strong objection to the approach comes from Minsky [9]; he says that the concept of the "baby machine" is reasonable, but that we do not yet have enough knowledge to build it. In particular, he states that "we do not yet have enough ideas about how to represent, organize, and use much of commonsense knowledge, let alone build a machine that could learn all of that automatically on its own". He also cites McCarthy to support this position: "in order for a program to be capable of learning something, it must first be able to represent that knowledge". Our argument above tackles this objection head-on and states that we must build a baby machine precisely because we do not know how to represent much of commonsense knowledge. AI has plenty of examples of programs that can go beyond the knowledge of their designers. Just as Samuel's checkers program played a better game than himself, so we could hope that a constructivist learner could find representations for knowledge which the designer does not know how to code. Obviously it will be challenging to build a learning program when we do not exactly know the target information to be acquired, but we do know the competence which should be displayed, and we know how the competence should change as a result of a certain experience. The designer will have to provide the initial knowledge and method of organisation; this will be refined in an iterative way during the development of a system to follow a particular development path, such as that outlined for the behaviour of the stick. A particular way of organising knowledge will be trialled, and doubtless shortcomings will be found when the system is not able to make the required knowledge acquisitions, and so the organisation will be tweaked, so that it can progress further along the development path, and so on.

A somewhat facile response McCarthy's statement above would be to state that there are learning techniques such as inductive logic programming or genetic programming which can represent pretty much anything, but this is probably not what he meant; they would take too long to find complex representations if starting without substantial background knowledge. In order to be able to learn something in a reasonable time, a program's existing background knowledge and representational framework should be at least pretty close to what is required for the new knowledge. The Piagetian approach proposes an incremental improvement of the representational framework. This is the main argument of the constructivist, which may be at odds with McCarthy; it is that you do not need to have a representational framework in place which is ready represent all required knowledge; frameworks can be tweaked to accommodate the requirements of new knowledge. An argument can be made by juxtaposing a modern scientist with a human adult from a hunter-gatherer tribe. The scientist has representational frameworks which are not present in the hunter, and the hunter could not immediately learn to apply some new algebraic formula. Yet, after an appropriate educational path is followed, the hunter could acquire the appropriate framework and learn the same knowledge. The point is that the hunter has the required constructivist machinery, and that is sufficient. Minsky *et al.* note that "You cannot teach algebra to a cat" [9], and indeed you cannot teach algebra to a baby either, but as a human continues to acquire more sophisticated background knowledge, then at some point it reaches a stage where it has the necessary scaffold on top of which algebraic knowledge can be constructed. By modelling infant acquisitions we would hope to find a computational model of the infant's learning machinery, including how to organise and use new knowledge, i.e. those mechanisms which a cat is lacking.

3.3 Summarising the Prospects

To conclude this section we will summarise the prospects for the constructivist approach to AI, and the promise it holds. By focussing on recreating a particular development path, we are forcing ourselves to come up with a mechanism which can add to its knowledge autonomously. This forces us to tackle the mechanism of intelligence as a central issue. An alternative (and more common) approach to AI is to try to recreate a particular behaviour or competence; the danger here is that specialist knowledge is coded in to solve that task,

and the the resulting system does not shed any light on general intelligence. The constructivist approach holds the promise of acquiring an understanding of concepts (such as size and weight), and having them grounded in sensorimotor behaviours. Such concepts could then serve as the base on which further concepts would be learned through language.

4 PROGRESS

Despite a sizable body of theory from Piaget and others, there has been relatively little work on constructivism in AI to date, and much remains to be tried out. For example, no AI work to date has attempted to model the infant's construction of space which has been sketched in Section 2.1. We will now review some of the main investigations which have been done in constructivist AI.

An ambitious attempt to model Piaget's description of the acquisitions during infancy is the doctoral thesis of Gary Drescher [7] . Drescher built a program to mimic the mechanism of early Piagetian development, and the way in which the concept of an object is learnt. Drescher's program worked in a 7x7 grid world, with a hand, eye and mouth. The program learnt "schemas" which consisted of a context, an action and a result. For example it learnt that if its current context was "HandInFrontOfMouth", and it took the action "HandBackwards", then it would expect to obtain the result "HandTouchingMouth". After exploration it was able to reliably predict the effects of most of its actions from whatever context it was in. Drescher also included a pair of objects in the world, which the eye could see, and the hand could touch and grab, etc. However, one object moved occasionally of its own accord, thus the schema for grabbing it was not entirely reliable. Drescher introduced the idea of the "synthetic item" to cope with this, the synthetic item could be "on" if grabbing had worked recently, and thus could be used to predict if grabbing the object was likely to work. The "synthetic item" is interesting because the program is starting to learn higher order data item which goes beyond what is directly sensed. This is in effect an abstraction of the raw sensor data which allows predictions to be made about objects in the world, and so it arrives at an extensional approximation of the concept of an object (i.e. it is generally "on" when the object is present).

Chaput's doctoral thesis [3] recreated the achievements of Drescher, and went significantly further. Chaput developed a "Constructivist Learning Architecture" which is based on Leslie Cohen's theory of infant cognitive development [4]; this is a neo-Piagetian theory which provides a little more detail than Piaget did about the required information processing mechanism. Cohen has abstracted, from many studies on infants, a set of information processing principles which apply throughout development and across all domains. These principles state that infants learn to process information at increasingly higher levels of abstraction by forming higher level units out of relationships among lower level units. There is a bias to process information using the highest formed units, unless the input becomes too complex, in which case the infant drops back to a lower level and attempts to refine its abstraction so as to be able to handle the complex information at the higher level. Essentially Cohen's principles describe a strategy for making abstractions; from the masses of raw data the infant need only pay attention to the abstractions it has found useful. Chaput's computational model successfully models some aspects of infant development, in particular the perception of causality. He then applied it to Drescher's microworld, and to a robot learning task. In the robot learning task the robot had to "forage" (i.e. see objects, move towards them, and pick them up). The robot had a vision system which had a 60° viewing angle, separated into five sectors of 12° each. When a blob appeared in one of these sectors, then a binary sensor item was set to true. Chaput used Cohen's information processing principles to construct synthetic items in an efficient way. In the foraging task, Chaput's system came up with many interesting and very useful synthetic items, which led to very efficient performance on the task. To compare Chaput's system with the system that would be required to learn the behaviour of the stick above, many of the required elements are there, but more machinery would be required to be able to make analogies between schemas. Chaput's robot sensors are too simple to allow such analogical machinery to be useful (the robot could not see one object in front of another for example).

Cohen et al. [6] provide a quite different approach to coding Piagetian schemas, which is somewhat more complicated, with action schemas containing "maps". These maps can represent a space with dimensions of distance and velocity, for example, and they can record the activation of a schema as a trajectory in this space. Their system can learn "gists" which are compositions of action schemas for certain tasks. This has been successfully applied to learn behaviours in a simulated world, for example a creature learns to sneak up on, and catch, a cat. The schemas learnt have also been transferred to similar situations in slightly different scenarios. This mechanism might well be applied to the developmental path we have outlined, in this case we would like to add to the mechanism so that it could see analogies between similar schemas or similar gists. This would allow a partial match to be found between two gists, and a conjecture to be made that unmatched aspects might also be similar.

A final related work which is worth citing is by Buisson [2]. This work learns to recognise the rhythm of a piece of music. The system features an active type of perception which attempts to match the rhythm being played by playing its own schemas to synchronise with it. It uses an evolutionary algorithm which starts with its own basic schema and generates mutated schemas, some of which will match with the target rhythm, and some not. Those which match will replicate and those which do not will die. Buisson's work seems to be the only computational investigation of Piaget's theory which really takes Piaget's idea of "assimilation" seriously. By assimilation Piaget means the way that experiences in the environment can be matched to known schemas. Buisson takes this seriously by acknowledging that the rhythm which is being played cannot be simply copied if it is not known already; the program must be active in conjecturing a variation on rhythms it knows to see if this might match the rhythm being played.

5 CHALLENGES

We will now outline some of the major challenges that need to be overcome to successfully exploit constructivist theories in AI.

(1) We need to find plausible development paths to copy. Piaget gives us some sketchy accounts, and much work will be required to flesh out the details of these. Research on infants could be particularly helpful here, in particular the type of study conducted by Bower [1], as discussed in Section 3 above. For example, we could investigate what kind of training would accelerate an infant's development of the behaviour of the stick, this would allow us to identify the experiences which infants profit from to develop the behaviour. We could also investigate what types of landmarks are being used by the infant, by investigating if their absence affects the behaviour. Such studies are very resource intensive unfortunately, but the idea that we could really settle some of these questions with infants is exciting.

(2) Given a particular development path, we need to find a computational model to explain it. Particularly tricky issues here include developing appropriate analogy finding mechanisms, finding appropriate representations for schemas of knowledge, and organisation among schemas (in particular sub- and super-schemas).

This challenge may not even be achievable in an AI system. Finding a development path in infancy would show that it is possible in the case of the infant, but whether this extends to an AI system (simulated or embodied robot) with different sensors and effectors is not clear, and will need to be investigated.

Assuming that it is possible, we can foresee the following type of iterative progress: the system will successfully reach a certain stage in the target development path, and will be incapable of making the next step. It will then be necessary to add innate knowledge or abilities to the initial infant, to see if that can allow it to go further. This is to be expected to some extent, because there may be some aspects of the infant's innate mechanism which have been present since the beginning, but just had no chance to express themselves before a certain level of knowledge was reached. Thus our system could reach a simpler level of knowledge without these abilities and then find that something is missing. The danger however is that, in helping the artificial infant to get to a particular milestone, we may put in innate abilities which could in fact have developed by themselves.

There are a number of examples in Piaget's theory of "new" behaviours that are expressed at a certain age, which seem to appear from nowhere; however, a detailed reading of his theory typically gives an account of how they are developed by the same mechanisms that have been at work since birth. An example of this is experimentation; the infant seems to suddenly start experimenting with the parameters of actions, and varying them, at about twelve months. However Piaget accounts for this by explaining that it arises because he now has so many schemas to recognise effects of actions; if a slightly different result is accidentally produced by an action, this difference will be salient to the infant, and the infant will try to ferret out the parameters of the action which can cause it. Thus the development of schemas which has been happening all along can account for the "sudden" emergence of experimentation.

It is also important that the world in which the artificial baby develops is sufficiently rich to allow the baby to develop all the behaviours we require. Otherwise it might fail to achieve some milestone not because of any deficiency in the innate mechanism. Getting this right requires a deep understanding of the development path we are trying to follow, and the nature of the knowledge which should be acquired; it will also doubtless require some trial and error.

(2.1) The challenge is really to explain the development with a minimal mechanism. We could say that the challenge of the programme as a whole is to show how qualitatively different behaviours can emerge from the continuous operation of a single mechanism (we are proposing that this mechanism be researched by trial and error with AI systems). We need to be careful about adding a new ability to the innate mechanism, when it might be possible to make the new ability emerge by providing a sufficiently rich world, or an appropriate developmental path which allows the existing mechanism to develop it.

(4) As the number of schemas grows, combinatorial problems will arise. It will not be possible to search for correlations between all schemas in order to find relationships to explain new phenomena; schemas will need to be searched selectively, and to be organised in some fashion. This is one of the problems that Minsky cited in his criticism of the baby approach. Possible solutions might be found in Cohen's [4] principles.

6 CONCLUSION

We have claimed that there is much value to be gained by applying constructivist theories to Artificial Intelligence. We advocated a research methodology which would set the modelling of a developmental path as a goal, rather than the achievement of some particular competence. This would force the research to investigate the constructivist mechanism itself (and hence analogy, among other things). It is hoped that this would lead to systems which are more adept at general tasks, when compared with classical AI systems (which achieve competence on constrained tasks).

We outlined a particular development path which might be attempted by the constructivist AI approach. This path followed the infant's laborious construction of space by analogy, and grounded in known actions. Because this construction is a laborious process, it must bring some benefit to be considered worthwhile. To justify this we explained that a construction which is built in this laborious way is very useful for a system that needs to act in the world, because its perception of the world is now built from more primitive actions which it knows. With this argument we hoped to point out the value of Piaget's theory, by showing why AI might do well to model this somewhat idiosyncratic development path.

We also saw that the main element currently missing from related work is the ability to find analogies among existing knowledge, and so to conjecture more elaborate models of the world. In this respect there is much more to be exploited in Piaget's theories.

Acknowledgements: Special thanks to Nir Oren for comments.

REFERENCES

[1] T. G. R. Bower, *Development in Infancy*, San Francisco : W.H. Freeman, 1982.

[2] Jean-Christophe Buisson, 'A rhythm recognition computer program to advocate interactivist perception ', *Cognitive Science*, **28**(1), 75–87, (February 2004).

[3] Harold Henry Chaput, *The constructivist learning architecture: a model of cognitive development for robust autonomous robots*, Ph.D. dissertation, Artificial Intelligence Laboratory, The University of Texas at Austin, 2004. Supervisors: Benjamin J. Kuipers and Risto Miikkulainen.

[4] L. B. Cohen, 'An information-processing approach to infant perception and cognition', in *The Development of Sensory, Motor, and Cognitive Capacities in Early Infancy*, eds., F. Simion and G. Butterworth, 277–300, East Sussex: Psychology Press, (1998).

[5] L. B. Cohen and C. H. Cashon, 'Infant perception and cognition', in *Comprehensive handbook of psychology. Volume 6, Developmental Psychology. II. Infancy*, eds., R. Lerner, A. Easterbrooks, and J. Mistry, 65–89, New York: Wiley and Sons, (2003).

[6] Paul R. Cohen, Yu-Han Chang, Clayton T. Morrison, and Carole R. Beal, 'Learning and transferring action schemas', in *IJCAI*, ed., Manuela M. Veloso, pp. 720–725, (2007).

[7] G. L. Drescher, *Made-Up Minds, A Constructivist Approach to Artificial Intelligence*, MIT Press, 1991.

[8] M. M. Haith, 'Who put the cog in infant cognition: Is rich interpretation too costly?', *Infant Behavior and Development*, **21**, 167–179, (1998).

[9] Marvin Minsky, Push Singh, and Aaron Sloman, 'The St. Thomas Common Sense Symposium: Designing Architectures for Human-Level Intelligence', *AI Magazine*, **25**(2), 113–124, (2004).

[10] J. Piaget, *The Origin of Intelligence in the Child.*, London: Routledge & Kegan Paul, 1936.

[11] J. Piaget, *Play, dreams and imitation in childhood*, London: Heinemann, 1945.

[12] J. Piaget, *Construction of reality in the child*, London: Routledge & Kegan Paul, 1954.

[13] M. Ridley, *Evolution*, Oxford: Blackwell, 2003.

[14] A. Sloman, J. Wyatt, N. Hawes, J. Chappell, and G. J. M. Kruijff, 'Long Term Requirements for Cognitive Robotics', *Proc. Cognitive Robotics '06 Workshop, AAAI'06*, (2006).