The Importance of Inferring Intention

Rodrigo da Silva Guerra
guerra@er.ams.eng.osaka-u.ac.jp

***Abstract*— This report compiles several aspects of cognitive studies collected through a multidisciplinary review of current ongoing research from different fields of study on the search for missing aspects on developmental robotics approach towards the understanding of human intelligence. Besides being far from complete, this review clearly leads to the unique *human capacity of understanding observed actions as goal-directed* as the one of the most fundamental *terra incognita* aspects of human nature, lying on the boundary between pragmatic and theoretical studies on modeling the development of human cognition. Though being the subject of previous studies in the light of classical AI approaches, strongly biased by strictly symbolic representations, the *intention inferencing process* still lacks deeper exploration from a modern embodied and developmental perspective.**

I. Introduction

Genetic evolution is known as being a very slow process. Evidences suggest that hominid evolution begun as early as 70 million years ago, and *homo sapiens* just like us were already on Earth about 50 thousand years ago [1]. Nevertheless, it is beyond any doubt that nowadays man differs from its first ancestors in many ways that go beyond physical mutation. After some stage of the evolutionary process, somehow, humans became able to evolve despite the need of significant physical changes in their bodies, which would cost thousands of years of natural selection. For instance, a hundred of years means almost nothing for biological evolution while on the other hand, just a few decades may represent a whole change in the development of our technology. Thus it is clear that biological evolution does not provide us with adequate explanatory power to elucidate recent human evolution as a whole. The time has come to ask what makes our culture so powerful, and what is so special about humans not merely looking to the individual only, but also to the cultural phenomena as a whole. Some authors have already warned about the importance of embodiment for the understanding of intelligence [2]. For reasons that are going to be developed in the next sections, I'm about to claim that not only embodiment, but also *population* aspects of human interaction that enable, in a broad sense, *human exchanging of ideas* should also be considered for a best approach on the way towards the fully understanding of human intelligence. Finally, based on that, this work will explore the importance of our early capacity of *assigning intention* to observed actions as a fundamental field of research for modeling human intelligence without missing global views from developmental, embodied and populational perspectives.

II. Cultural Evolution

It has already been said that biological evolution alone is not sufficiently adequate to explain recent human evolution. For instance, Ms. Gabora wisely illustrates this deficiency of natural selection as an explanation for the current stage of evolution with a comparison [3]. As she says, if we were to go back some time in the first billion years of Earth's history, the only causal principle we would need to invoke to explain pattern in information would be the physical constraints and self-organizing properties of matter. But some time after the origin of life, let's say, about three billion years ago, this would no longer be the case because it would be impossible, in her example, for a giraffe to appear in an information space not acted upon by natural selection. There was another causal principle to be invoked from that point on. Today the Earth is embedded with artifacts whose existence cannot be explained neither by the properties of matter nor by biological evolution. As she says, biological evolution is not sufficient to explain the existence of computers just like the self-organizing properties of matter are not enough to explain the existence of giraffes. For instance, cars and computers are manifestations of yet another causal principle: the evolution of culture [3]. She thus classifies the patterns encountered in the structure and dynamics of information into three broad causal principles:

1) The physical constraints and self-organizing properties of matter;
2) Biological evolution;
3) Cultural evolution.

Having said that, it turns out to sound reasonable having some more formal understanding of the underlying processes that allow the development of such a recent and powerful causal principle.

III. Memetics

Several authors point towards the possibility of looking at cultural development as a second form of evolution, through the perspective of *memetics* [4], [5], [6], [7]. Memetics model cultural evolution, including the evolution of knowledge, through the same basic principles of variation and selection that underlie genetic evolution, which rely on the existence of [3]:

1) A pattern of information;
2) A way to generate variations of the pattern;
3) A rationale for selecting variations that tend to give better performance than their predecessors in the context of some problem or set of constraints;
4) A way of replicating and transmitting the selected variations.

The straightforward implication of these assumptions is the existence of a cultural replicator, basic unit of cultural information, named *meme*, the memetic counterpart of the *gene* [4]. Thus the meme could be understood as the common atomic unit from which all sort of human cultural expression is made of. The Compact Oxford English Dictionary defines a meme as "an element of behavior or culture passed on by imitation or other non-genetic means" [8]. A meme, just like the gene, is an arbitrarily defined concept, but with not so straightforward link to any special physical pattern as genes do. And unlike genes, memes do not come with instructions for their reproduction. They rely on their hosts to create, select, and replicate them accordingly to the physical, biological and cultural needs they satisfy.

The power of this approach lies in concentrating on the *meme* as the object of the evolutionary process that makes cognition possible, rather than concentrating on the individual. This perspective discloses population level phenomena that would not be readily detected through introspection. Memetics provides thus a different synthetic framework for understanding how mental representations are generated [3].

But while, in one hand, for sociology, economy, politics and other fields it may be sufficient the understanding of population aspects of the cultural development itself, on the other hand it is of special interest for psychology, cognitive sciences and developmental robotics the understanding of the means by which human culture turns out to be possible and the basic cognitive aspects that allow this most powerful and unique human characteristic to raise. The power of the synthetic approach of developmental robotics should account for the aspects which turn out to be essential for the emergence of cultural development in an artificial system thus giving more complete insights about the nature of human intelligence.

The next few sections are intended to show from which point of view it seems to be clear that more special attention should be given to the development of the so-called *theory of mind* as a crucial step on the development of human culture. Developing a theory of mind allows the child to understand other people, and to predict what other people are likely to do and believe. It has been already suggested that theory of mind should be seen as an emergent behavior that has some basis on humans early capacity of understanding other agents' actions as goal-directed ones. Based on this I will suggest the that advantages may rise by taking the developmental robotics embodied approach on the study of our capacity of *understanding of other's intention.* The embodiment perspective may give powerful insights on the rising of *humans inferring capacity* on the interaction of one agent with other agents and its environment.

IV. Social Learning and Imitation

Looking at culture as an evolutionary system through the perspective of memetics requires the acquaintance of the hypothesis of existence of a replicator, which in this case is conceived as the meme. From the exchanging of many of these memes among human society the culture would be supposed to rise and evolve through a selective process. Besides existing a certain degree of ambiguousness on the definition of meme, most authors agree that the definition of meme depends, if only in a broad sense, on the concept of imitation. For instance, Ms. Blackmore claims that one should consider *true imitation* alone as the only memetic process, excluding from consideration all other kinds of social learning [9]. Her argument was a source of divergence because imitation is just one of several processes that can result in social learning. Actually in the majority of cases where the transmission process is investigated, behavior patterns are not transmitted by imitation, but instead result from other, simpler processes such as local enhancement, where an animal's attention is drawn to an object by actions of another, in a manner that it results in learning [10]. Besides these controversies there is a general agreement about the importance of the unique human imitation of new behavior patterns to the evolutionary power of the human culture. As Ms. Blackmore claims, humans are capable of extensive and generalized imitation giving rise to new replicators that can propagate from man to man, or from man to artifact and again to man, thus making culture possible [9].

Several evidences suggest that our complex language skills evolve from infancy relying on innate mechanisms. Nevertheless it is still due to our imitation capacity that we acquire language practical abilities ranging from the reproduction of simple phonemes up to the most complex aspects of syntax. As it may be suggested by reading next section, there is a possibility that the most fundamental aspects of language, such as the semantic grounding, the simple gestural interpretation and even the most basic conjectures humans can do about the environment are probably strongly related to our (perhaps innate) capacity of *projecting intention* on observed actions. Finally, the human ability of developing the *theory of mind* is perhaps the most important aspect, unique to humans, which may rise as a conceivable implication of humans early inferential abilities, possibly evolving facilitated by our imitation skills, probably enhanced by innate mechanisms such as, for instance, the *mirror neurons.*

V. Goal Directed Actions

Preverbal infants as young as 1 year old already tend to interpret other people's actions as goal-directed, extending these interpretations to the behavior of other animals and even to inanimate moving objects. It has been already proposed that early imitation of goal-directed actions in infants is a selective, inferential process that involves evaluation of the rationality of the means in relation to the constraints of the situation - known as *the principle of rational action.* Infants expect not to see unnecessary, unjustified actions in observed behaviors, and when they do see such actions, they don't consider those as means to the goal [11]. Infants can separate the goal from the means, and can choose to use the same newly demonstrated action if they consider it to be the most rational alternative to achieve that same goal [12]. Such imitative learning is thought to be specific to humans, as primates do not imitate new strategies to achieve goals, relying instead on motor actions already in their repertoire (emulation).

More than that, one-year-old infants already use the principle of rational action not only for the interpretation and prediction of goal-directed actions, but also to make productive inferences about unseen aspects of their context [11]. Several theorists proposed that this precocious ability reflects the first

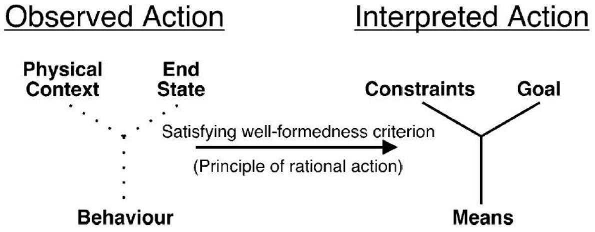

Fig. 1. Teleological representation of goal-directed actions (extracted from [11]).

step towards the development of the *theory of mind*, that is, the ability to understand others as intentional agents. In the following subsections some experiments with infants which lead to these conclusions are summarized together with a recently proposed approach on the representation of goal-directed actions.

A. Teleological representation of goal-directed actions

Gergely and Csibra proposed for a goal-directed action to be represented in terms of a so-called *teleological interpretational schema* containing three elements: (a) the observed behavior, (b) a possible future state, and (c) the relevant aspects of physical reality that constrain possible actions. By definition, as shown in figure 1, a *well-formed* teleological representation *only* happens when the observed behavior can be considered as an effective way towards the future state given the physical constraints of the situation. If this condition is satisfied, the future state will be interpreted as the goal, the behavior as means to this goal and the relevant aspects of physical reality as action constraints.

In a recent invited paper for the epigenetic robotic community, Mr. Gergely summarizes some recent research findings that came about while experiencing with pre-verbal infants [13]. By referring those experiments, Mr Gergely conjectures about the theory of *teleological representation* as being a possible framework from which infants rely for explaining relation between the three main aspects (see figure 1) of the observed behaviors. All those experiments were performed with basis on the violation-of-expectation looking time, which basically is founded on the premise that infants probably spend more time giving attention to a scene when this scene shows some unexpected behavior. A short summary of those experiments is presented in the following paragraphs.

a) Experiment A: In this experiment one-year-old infants were habituated to a computer animated goal-directed action (figure 2 A) in which a small circle repeatedly approached and contacted a large circle (goal) by jumping over (means act) an obstacle separating them (situational constraint). To test if whether or not one-year-olds would also interpret this event as goal-directed, it was presented to them two another test events in which the obstacle was no longer present. The first of those test events generated an incompatible outcome in which the already familiar jumping approach is used besides being no longer justified as a rational action, because the obstacle had been removed. The second of those test events generated an compatible outcome, but perceptually new, a coherent straight-line goal-approach, which had became a more reasonable action since the obstacle was removed.

Results showed that in spite of the fact that the old jumping approach was perceptually similar to the means action presented during the habituation event, subjects looked at it significantly longer than at the novel straight-line approach. Even though the straight-line approach was perceptually novel, infants showed no dishabituation to it.

b) Experiment B: In this second experiment, during habituation, one-year-old infants observed a large ball which was approaching a moving small ball until the latter passed through a small gap between two obstacles and disappeared from the screen (figure 2 B). The larger ball, being too big to pass through the gap, had to make a detour around the obstacles, in order to continue supposedly behavior chasing the small ball, until also disappearing from view. After the habituation, two test events were presented, both of which the upper part of the screen was opened up, revealing before unseen future states. In the first of the test events an incompatible outcome is shown, in which the small ball stops and the larger ball simply passes by it, without ever catching the small ball. In the second test event the small ball also stops in the now visible upper part of the screen, but this time the larger ball showed a congruent behavior by successfully chasing the small ball, stopping over the position in which the small ball was posed.

Infants looked at the incongruent test display for a significantly longer period of time than to the congruent test display, suggesting that the incongruent outcome violated their expectation about the goal state they had inferred as a rational completion for the unseen future states of the control event, characterizing it as a successful chasing.

c) Experiment C: In this third experiment, during the control event, a square shaped plate is placed so as to occlude from the child's vision significative portion of the place of action, while a small ball comes from the left. This small ball thus moves to the right until it becomes partially hidden by the plate, then showing a partially occluded yet perceptible jump, thus further proceeding its movement up to the encounter of a bigger ball placed in the extreme left side (figure 2 C). The two test events consisted of different versions of the same action performed without the hiding plate. In the first, incongruent test event the plate is lifted showing no obstacle for which it would be possible to justify the jumping of the ball. In the second test event the plate un-hides an obstacle over which the ball justifiably jumps and follows on its way towards the larger ball on the left.

One-year-old infants again looked significantly longer at the incongruent display than at the incongruent one, indicating that they inferred the presence of some occluded obstacle in order to justify the jumping approach as an efficient means to the goal.

B. Inferential principle of rational action

These results indicated that by 12 months infants can:

1) Interpret an other agent's action as goal-directed;

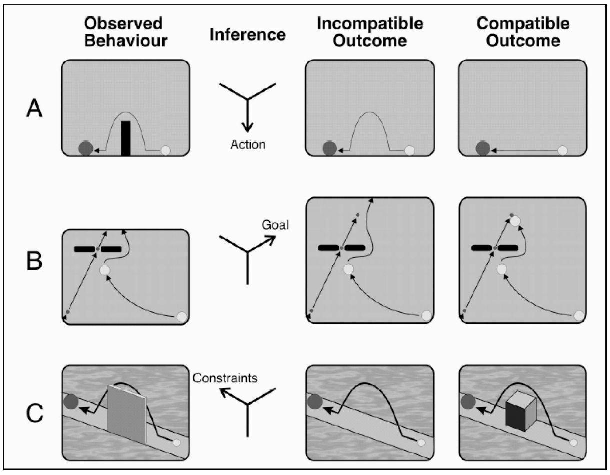

Fig. 2. Inference of teleological missing elements (extracted from [13]).

2) Evaluate which one of alternative actions available within the constraints of the situation is the most efficient means to the goal, and;
3) Expect the agent to perform the most efficient means available in the given situation to pursue the goal.

By showing these experiments, Mr. Gergely argues that one-year-olds can go beyond the information perceptually given and can productively infer any one of the three representational aspects of the teleological representation of a goal-directed action, which are (a) means act, (b) goal state, and (c) relevant situational constraints. If any one of these aspects is perceptually inaccessible to the infants, this aspect will be inferred, as long as the infants have direct perceptual information about the contents of the other two elements. To interpret the experiment actions as efficient and justifiable goal-approaches, the infants had to use the rationality principle to infer and fill in the content of the relevant missing element of the representation [13].

C. Inferential imitation

Infants imitative mechanism is functionally selective, and this should be accounted by the inferential capacities that constrain and guide infants behavior [13]. In the experiment described bellow infants show their ability to learn new means by imitation from observing adult's demonstration:

d) Light box experiment: The infants observed the adult turning on a light, illuminating a box, by touching its top panel with the forehead. In one condition the infants could see that the demonstrator's hands were occupied while demonstrating the executed action. Pretending to be cold, the adult had wrapped a blanket around herself which she held onto with both hands, showing the hands as being clearly occupied (see figure 3 A). In another condition the same experiment was performed, but this time the adults hand where free, thus not justifying the forehead approach in this case (see figure 3 B) [12].

As a result, 69% of the infants who observed the hands-free experiment imitated the same head action. However, after watching the adult turn on the light with the head while the hands where occupied, the number of children who imitated the head action dropped to 21%. This shows that infants were sensible to the fact that the demonstration should use head action when the hands were occupied. When observing the demonstrator hands were occupied, children chose not to imitate the head action presumably concluding that the head action was not the most rational [12].

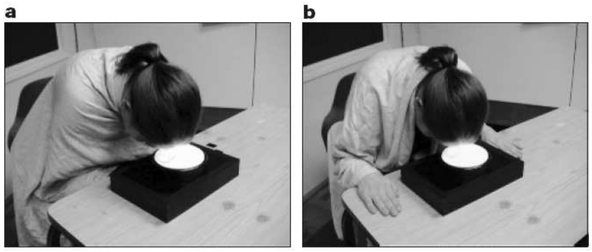

Fig. 3. Light-box experiment (extracted from [12]).

Fig. 4. Ipsi-lateral imitation error (extracted from [13]).

e) Ipsi-lateral imitative error.: In this experiment children between 3 and 6 years of age were asked to imitate an adult model's goal directed target actions that involved touching either of their own left or right ear with either a ipsi-lateral (i.e. right hand on right ear) or a contra-lateral (i.e. right hand on left ear) hand movement (figure).

They found that the children were practically errorless in reproducing the goal of the demonstrated actions, always touching their correct ear, corresponded to that of the adult's demonstration. But, nevertheless, children were susceptible to commit a specific type of error: when contra-lateral hand actions were demonstrated to them, for example, when the adult touched his left ear with his right hand reaching across his body, the infants very often touched their corresponding ear with a ipsi-lateral rather than a contra-lateral hand movement. Children in this case failed in imitating the specific means of action. In their attempt to realize the same goal state as the adult, they tended to use a simpler, more familiar, and thus more rational alternative means.

VI. Conclusion

Let's dicuss about possible conclusions.

References

[1] R. Larick and R. L. Ciochon, "The african emergence and early asian dispersals of the genus homo," *American Scientist*, pp. 538–551, November, December 1996.

[2] R. Pfeifer and C. Scheier, *Understanding Intelligence*. The MIT Press, 1999.

[3] L. Gabora, "A day in the life of a meme," *Philosophica*, no. 57, pp. 901–938, 1996.

[4] R. Dawkins, *The Selfish Gene*. Oxford University Press, 1976.

[5] Cavalli-Sforza and L. Luca, *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, May 1981.

[6] C. Lumsden and E. O. Wilson, *Genes, Mind, and Culture*. Harvard University Press, 1981.

[7] P. Schuster and K. Sigmund, "Replicator dynamics," *Journal of Theoretical Biology*, no. 100, 1983.

[8] C. Soanes, ed., *Compact Oxford English Dictionary of Current English*. Oxford University Press, second ed., June 2003.

[9] S. Blackmore, "Imitation and the definition of a meme," *Journal of Memetics - Evolutionary Models of Information Transmission*, vol. 2, pp. 159–170, December 1998.

[10] S. M. Reader and K. N. Laland, "Do animals have memes?," *Journal of Memetics - Evolutionary Models of Information Transmission*, vol. 3, no. 2, pp. 100–108, 1999.

[11] G. Csibra, S. Biro, O. Koos, and G. Gergely, "One-year-old infants use teleological representations of actions productively," *Cognitive Science*, no. 27, pp. 111–133, 2003.

[12] G. Gergely, H. Bekkering, and I. Kiraly, "Rational imitation in preverbal infants," *Nature*, vol. 415, p. 755, February 2002.

[13] G. Gergely, "What should a robot learn from an infant?," in *Proceedings of the Third International Workshop on Epigenetic Robotics* (C. G. Prince, L. Berthouze, H. Kozima, D. Bullock, G. Stojanov, and C. Balkenius, eds.), (Boston, USA), pp. 13–24, Lund University Cognitive Studies, August 2003.