# The Stored Interactions Model (SIM), an interactivist model of Piagetian assimilation − application to visual perception

**Jean−Christophe BUISSON**

IRIT (Computer Research Institute of Toulouse, France),
ENSEEIHT (Toulouse, France) buisson@enseeiht.fr

Paper presented at the 31st Annual Meeting of the Jean Piaget Society,
May 31−June 2, 2001, Berkley, California

## 1− Introduction

My purpose is to find a mechanism, implementable on a computer, of the Piagetian assimilation process. Gary Drescher (1991) at MIT made a similar attempt, with interesting results concerning the emergence of causality and the construction of the notion of object, but the underlying mechanism he used operated on a very small and simplified world, and was computationally intractable for larger ones.

This story is common in the field of artificial intelligence and robotics: most if not all projects failed the scaling move, when augmenting the complexity of the world they were operating on. One of my main goals is to show that, on the contrary, perception processes based on the agent's activity are the key to solving this scaling problem, and to deal with the complexity of the real world.

## 2− Main ideas of the model

**Idea 1−** An interaction is a rhythmic succession of actions, each step being assorted with a set of sensory anticipations.

This idea captures the active nature of perception, where an action is made essentially to test whether an anticipated sensation will occur. For instance, a breath in is made to test the presence of a particular smell. An ocular movement is made to test the presence of a sharp vertical edge at the center of the retina. These steps of action/confirmation are chained to differentiate situations of increasing complexity.

**Idea 2−** Each and every interaction of the agent with its environment is recorded. This recording does not necessarily amount to a large volume of data, since only the sensorial and motor elements actively implicated in the interaction are stored.

**Idea 3−** All past stored interactions try permanently to get synchronized spatially and temporally with the current situation, in order to influence the on−going interaction. In a word, the current interaction is guided by previous ones, and will itself serve later as a guide for others.

Ideas 2 and 3 will be clearer with a metaphor. Imagine a circus stage, where all spectators would be able to influence by voice the lion tamer actions. The tamer is the current interaction, and the spectators are the stored interactions. Suppose that at the beginning, there are no spectators, and a lion tamer completely novice. When faced with a particular lion, he acts almost randomly, and is quickly injured. He then becomes a spectator, and a new tamer is placed on the stage. This second tamer does not act as randomly as the first, since he is now guided by the first spectator, who intently watches the performance. This spectator is able to (unreliably) anticipate some aspects of the situation, and may influence one of the tamer's action in the right way. After a time however the second tamer is injured, becomes an influencing spectator, and a third tamer appears, etc. As the successive tamers are confronted to different situations, differentiations appear between spectators. Those who survived a tiger for a long period will best understand and assimilate a tiger situation, and so will influence the tamer more strongly than would spectators who were confronted with lions.

Most of it is there. But there are many ways to instantiate this general model into one which is detailed enough to allow a computer implementation, and be "tested" in the real world. Of what, precisely, is an interaction composed of? When does it start, or stop? How a past interaction may be synchronized with the current situation? Etc.

Nevertheless, general as they are, these ideas have several important consequences:

**a− Piagetian assimilation schemes**

Activity is at the heart of our model, and stored interactions are assimilation schemes more operationally specified.

More precisely, the equivalent of a Piagetian scheme is the class of stored interactions which are activated by similar situations. Like Piagetian schemes, they try to apply whenever possible.

For example, suppose you are performing a shot at tennis. Your movements are guided by the subset of stored interactions reacting to this particular situation, and this appropriation is what Piaget calls assimilation. Accomodation has two aspects. It is first the adjustment, the synchronization process by which a stored interaction gets finely tuned to the situation. If you do not pay enough attention, you will perform basically the right movements, but they will not be precisely enough adjusted to the ball trajectory, and will result in a bad shot. A different aspect of accomodation is the addition of this particular shot movement, which may enrich the experience storehouse with a new interesting aspect of its own.

**b− interactivism and epistemic contact**

According to interactivism (Bickhard, 1995), cognitive representation cannot solely be composed of correspondences (like symbolic encodings), since it would lack any kind of real epistemic contact with reality. Bickhard shows that a more basic form of representation is necessary, interactive in nature, and mainly composed of functional indicators differentiating internal states and guiding the agent actions.
He shows, 1− that such indicators account for at least a functional form of representation, 2− that interactive representations can emerge from non representational grounds, so there are no symbol grounding problem, 3− an interactive representation is not an encoding, but can serve as a ground for encodings.

Our stored interactions are interactive representations. By guiding the current interaction, they functionally differentiate the agent internal states. They are not encodings. They emerge from non representational grounds, so they provide a genuine epistemic contact.

**c− learning**

Ideas 2 and 3 form a framework for learning; a learning which is permanent and continuous. Idea 2 is rather extreme: all interactions are recorded. How could we be able to access every aspect of our past activities? First, a complete recording is not necessary, and there are ways to compress the data. And it is only the motor and sensorial aspects effectively used during an interaction which are to be stored. But even a complete recording is not completely absurd. There are proven cases of prodigious memory capabilities (A. R. Luria, 1988). And we all experience as we get older the recollection of past scenes we thought definitely lost. Computationally, to store billions of the small amount of data each interaction is composed of is not intractable. In fact, that is what our model predicts: the brain stores this past activity, in a way which gives it the ability to guide instantaneously and even anticipates our current actions and thoughts.

This learning is also unsupervised. The current interaction stops when it is no longer guided by stored interactions. It is then stored itself, and a new current interaction is spawned. The stopping of the current interaction is the kind of error signal described by Bickhard: it occurs when the system is functionally unprepared for the situation. It allows this interaction to become part of the past, and to be used for the future, and that is the "positive" face of learning. It also tells the agent that the current way of interacting must be changed, and this is the "negative" face of learning. It resembles what goes on when someone retires: her past activities are remembered and used to guide current activities, but at the same time the person who replaces her now acts in a new way.

**d− perception**

This model has important consequences for perception, so I will subclass them.

<u>Perception is active</u>

To smell, we need to breath in. To touch, we need to move our fingers. To see, we need the ocular movements, and recent evidence (O'Regan, 2000) revived the old idea that vision is a sense of touch at distance.
Idea 1 tries to capture these facts: when perceiving, we perform actions in order to test anticipated sensations.

**<u>When looking at an apple, we see all the apples we have seen in our lives</u>**

All stored interactions related to seeing an apple are activated by looking at the fruit, and guide the way we look at it.

**<u>I can sit on it, so it is a chair</u>**

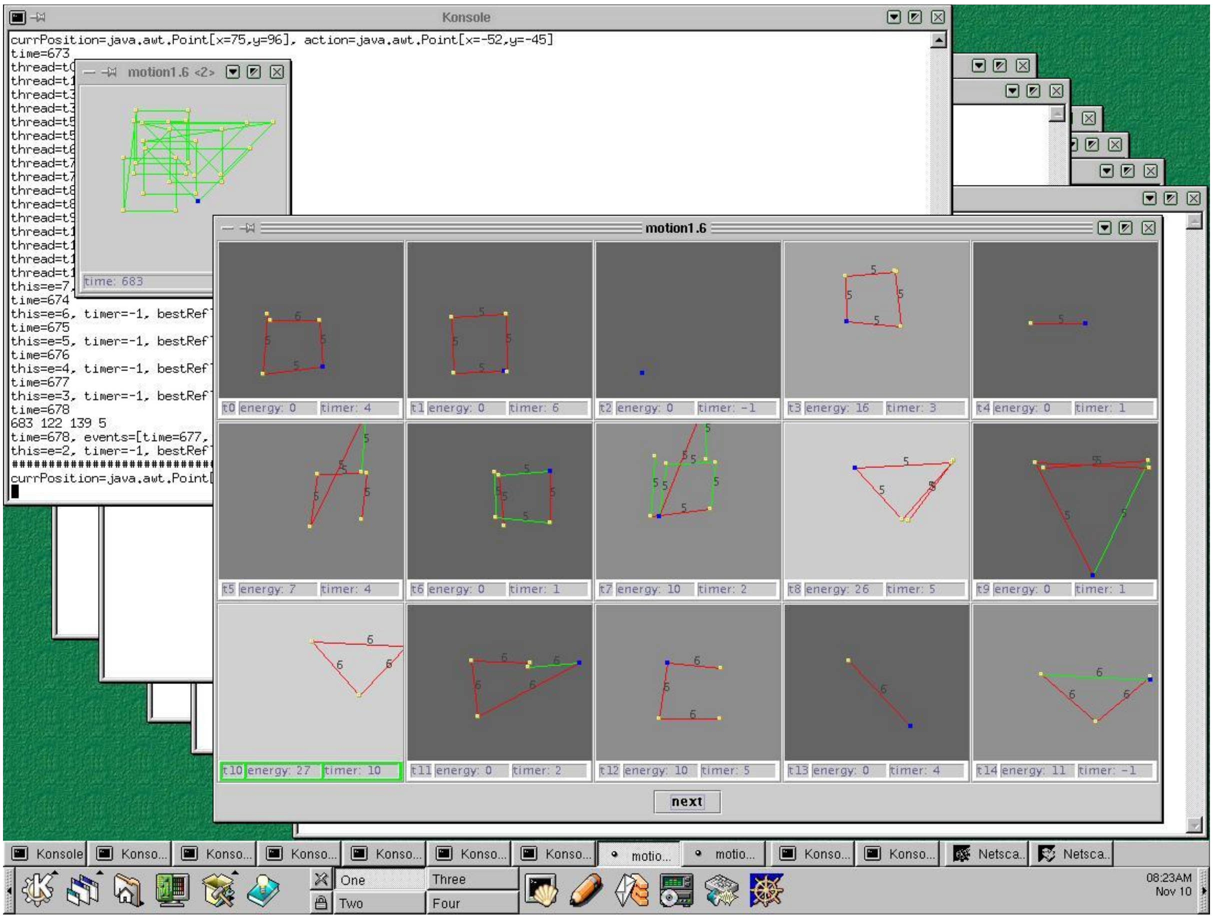
Most other mind theories would have it the other way round: a perception stage would discriminate objects on the scene, then one of these objects would be classified as a chair, and so would be used to sit on. R. Brooks (1991) for instance commented on the problems associated to each of these steps. They do not exist with interactivist approaches. Faced with a chair, even of a new kind, an agent will be urged by its stored interactions to engage in a sitting process, and this act, or the potentiality of this act is what can be called the chair recognition.

**<u>A way to deal with the complexity of the real perceptive world</u>**

As we said in the introduction, most A.I. approaches failed when moving from a toy world to a more complex one. I think that the solution of this problem lies in Idea 1, which expresses the fact that the agent deliberately tries to provoke anticipated sensations with its actions. Whatever the complexity of the sense flows, the agent will pick in the middle of them the pieces of information needed to confirm he is engaged on the right assimilation processes.

## 3− The simulation

Instead of explaining it abstractedly, I will present a screenshot of a running demo.




The nodes in the small window are the positions of past flashing lights. A random generator produces flashing geometric patterns (squares and triangles at present, randomly) with different flashing timings, sizes, positions and regularities. In the big window, we can see how the eye interacts with this scene in real time. Each tile in the window is a past interaction, the current one being the lower right tile. Edges represent ocular movements, and they are green when the action has been guided by a previous interaction, red when they are reflex. There are more and more green edges, as the system gets habituated to the patterns, and actively tracks anticipated events. Each previous interaction tries to assimilate the current situation by synchronizing on it, spatially and temporally; the degree of this assimilation is quantified in the 'activity' number, and reflected in the color of the corresponding tile. The best ones influence the course of the current interaction.

## Conclusion

I have presented an interactivist model of the Piagetian assimilation. It is also a learning framework of a new kind, which keeps intact each individual experience of the agent, and at the same time synthesize on the fly this knowledge to guide the current interaction. It contrasts with most learning algorithms based on neural networks, where past experience in melted into a small set of synaptic weights. Future experiments will show if the SIM model is computationally tractable, and if it can be extended to other domains.

## References

Bickhard, M. H., Terveen, L. (1995). *Foundational Issues in Artificial Intelligence and Cognitive Science. Impasse and Solution*. Amsterdam: Elsevier.

Brooks, R. A. (1991). Intelligence without Representation. *Artificial Intelligence*, 47(1−3), 139−159.

Drescher, G. L. (1991). *Made−Up Minds*. Cambridge: MIT Press.

Luria, A. R. (1988). *The Mind of a Mnemonist*. Cambridge: Harvard University Press.

O’Regan, J. K. (1992). Solving the "Real" Mysteries of Visual Perception: The World as an Outside Memory. *Canadian Journal of Psychology*, 46(3), 461−488.