AN AUTONOMOUS INTELLIGENT AGENT ARCHITECTURE BASED ON CONSTRUCTIVIST AI

Filipo Studzinski Perotto, Rosa Vicari and Luís Otávio Alvares
Universidade Federal do Rio Grande do Sul, Caixa Postal 15.064, CEP 91.501-970, Porto Alegre, RS, Brasil

Abstract: This paper intends to propose an learning agent architecture based on Constructivist AI approach. First we show how Artificial Intelligence can incorporate some concepts from Jean Piaget's psychology to form Constructivist AI as a branch of Symbolic AI. Then we present details about the proposed architecture, discussing the provided agent autonomy. Eventually, we report experimental results.

Key words: Artificial Intelligence, Constructivism, Agents.

1. INTRODUCTION

Over the last decade, Artificial Intelligence has seen the emergence of proposals of new paradigms, new techniques (many of them born from the hybridization of classical techniques), and a search for new theoretical conceptions, coming from the contact with other disciplines. Exploration on these AI borders may make possible to overcome some current limitations of conventional AI systems.

From these new ways of conceiving AI, we highlight the Constructivist Artificial Intelligence. In general lines it comprises all works on this science that refer to the Constructivist Psychological Theory, essentially represented by Jean Piaget. The Constructivist conception of intelligence was brought to the scientific field of Artificial Intelligence in the early 90's, and it has not occupied the desirable space among researchers, who are mostly linked to the classical paradigms. But even as an alternative proposal, still searching for legitimacy and consolidation, Constructivist AI has shown to be able of contributing to the discipline. Theoretical discussions on the meaning of the incorporation of piagetian concepts by AI are present in [Drescher 1991], [Boden 1979], [Inhelder & Cellerier 1992] and [Rocha Costa 1989, 1995].

Following the proposal of Constructivist AI, our work proposes a model of intelligent agent which implements some of Jean Piaget's-conceived intelligence and learning mechanisms. There are some models that are already implemented [Drescher 1991], [Wazlawick 1993] and [Muñoz 1999] which are our basic reference, both for the initial inspiration and as a parameter to see which limits need to be surmounted.

2. JEAN PIAGET'S PSYCHOLOGICAL THEORY

Piagetian Psychology is also known as Constructivist Theory. According to Piaget, the human being is born with a few cognitive structures, but this initial knowledge enables the subject to build new cognitive structures through the active interaction with the environment [Becker 2001], [Montangero 1997].

Piaget sees intelligence as adaptation, which is the external aspect of an internal process of organization of cognitive structures, called "schemas". So, the subject's schemas are transformed, but the functions that regulate the construction of new schemas do not vary during life [Piaget 1953]. A schema is all that, in an action, can be differentiated, generalized and applied to other situations [Montangero 1997]. It is the elementary cognitive structure, and represents a world's regularity observed by the subject. The schema comprises perception, action and expectation.

The subject realizes all of his experiences by his schemas. Sometimes the subject's expectation fails, and compels to a modication in schemas. There is, thus, a process of interactive regulation between the subject and the environment. The subject is all the time making "accommodations" and "assimilations". Assimilation is the process through which the subject uses his schemas to interact with the world. Accommodation occurs when his schemas are not able of responding to some situation, and is the process through which the subject modifies his schemas trying to adapt himself to the environment. In this game of assimilations and accommodations the subject progressively differentiates its schemas in order to deal with reality [Piaget 1978].

Accommodation is the transformation of schemas starting from the experience aiming at making the system more adapted to the environment. As accommodation happens in a schema the subject already has, the new structure arises molded by structures that already exist. Thus, at the same time that the schemas integrate novelties, they also maintain what they already know. The system reaches a stable point because each accommodation widens its ability of assimilating. Thus, novelties will less and less affect the system equilibrium, and the system will be more prepared to deal with novelties [Piaget 1953].

Initially, the psychological subject counts only on “sensorial boards” (instant perceptions of the visual, auditory, tactile, cinestesic, tasting fields ...). These boards do not maintain any implicit logic relation, and they are, a priori, disconnected. From this universe of “sensorial boards”, intelligence will build elementary notions, set relations, find regularities and eventually will build an objective, substantial, spatial, temporal, regular, external universe, independently of intelligence itself. The “real” will be composed in this adaptation movement by the increasing coherence between schemas [Piaget 1957].

3. CONSTRUCTIVIST AGENT: MODEL

The Constructivist Agent model we are proposing uses basically the concepts of Piaget’s Constructivist Theory. Previous models had already tried this accomplishment, each one using some type of computational technique to implement the mechanisms described by Piaget. The first model is presented in [Drescher 1991], and uses statistical calculation to find correlations between the agent’s observation, actions and results.

Other important model is presented in [Wazlawick 1993], and is inspired by genetic algorithms [Holland 1992] and in the model of self-organized neural network [Kohonen 1989]. In this architecture, the agent has a population of schemas, and each generation preserves the best schemas, that represents the correct relations between observation, action and results, based on the own agent experience.

Other Constructivist agent architectures are presented in [Muñoz 1999], that inserts planning to Wazlawick Agent, [Balpaeme, Steels & Looveren 1998], that uses the same strategy to reach emergence of catergories and concepts in a multiagent environment, [Birk & Paul 2000], which inserts genetic programming in Drescher Agent, and [De Jong 1999], using schemas that reinforces itself to form concepts.

The model we are proposing differs from previous ones because they are based only in assimilation-accomodation mechanisms, without statistical aproachs, or selectionist regulations, or reinforcement learnings.

An autonomous agent, in Artificial Intelligence, is a system able to choose its actions independently and effectively through its sensors and effectors in the environment. [Davidsson 1995]. The constructivist agent is an embodied agent. Agent's body acts as a mediator between environment and mind through its external sensors (inputs) and external effectors (outputs). Body also has internal (somatic) states and metabolisms that serve as basis to an internal organic motivational sense. Agent's mind interacts with body by internal sensors and internal effectors. The mind has two systems: cognitive and emotional. Figure 1 shows the agent's structure, with a body mediating the mind and the environment, and with an emotional system interacting with the cognitive system.

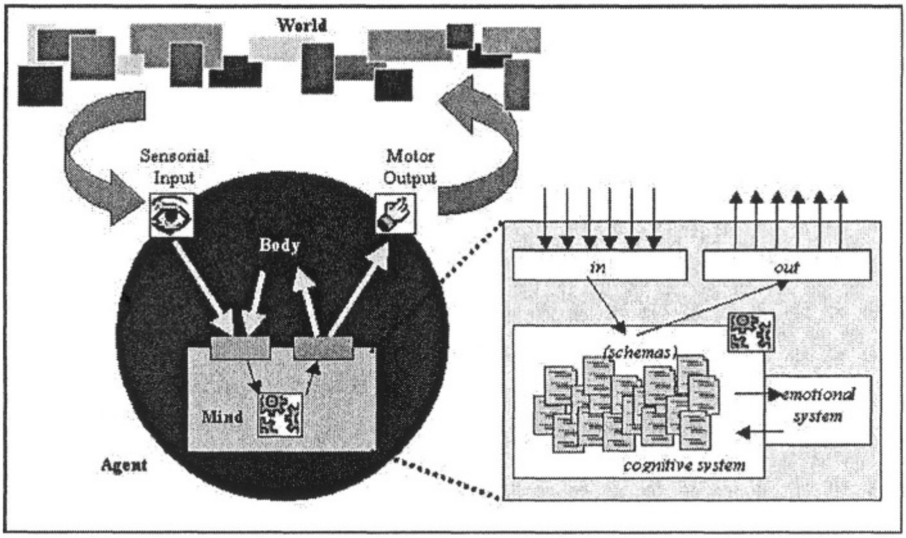

Figure 1. Architecture of Agent (on the left, the agent interacts with the environment through its body; on the right, details of mental structure)

The constructivist agent's mind receives sensorial inputs and activates action outputs. The cognitive system has a set of schemas (its cognitive constructions) representing agent's beliefs. For each situation the agent experiences, and depending on the desire of changing this situation, the agent will select and activate one of these schemas. Cognitive system also has a mechanism to build schemas. The set of agent's schemas may be initially empty because the mechanism proposed is able to build all its knowledge by interacting with the environment, while it carries out its activity not needing any pre-programming.

A schema is composed of a triple {Context, Action, Expectation}. The *Context* is the representation of situations that the schema is able to assimilate. *Action* represents the action that the agent will carry out in the environment if the schema is activated. *Expectation* represents the result expected after the action. A schema is represented in Figure 2.

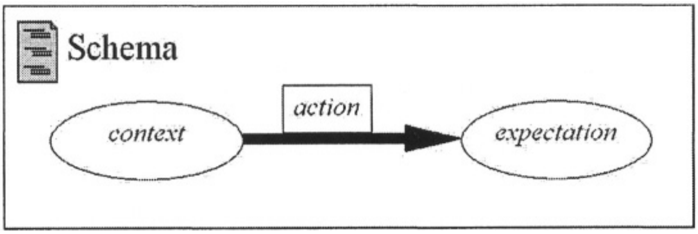

Figure 2. Structure of Schema (Context, Action, Expectation)

The situation in which the agent is at every moment (context realized through sensors) is compared with the context of the schema, and it will be excited if there is compatibility. When an excited schema is activated, the cognitive system makes it to perform its Action (through the agent's effectors). Regulation happens because the cognitive mechanism checks if the result (context realized at the instant following the action) is according to the expectation of the schema, where the difference serves as a parameter for adjusting and evaluating the schema.

At every instant (execution step), the agent's cognitive system verifies the context of all schemas, exciting those that are compatible with the situation realized through sensors. The mechanism consults, then, the emotional system, which will choose among the schemas excited the one that will be activated.

The mind's emotional system interacts with cognitive system and it is the responsible for guiding the agent's actions, conferring an affective meaning to knowledge and enabling the agent to have its own goals. Emotions can be seen as the subject's presence of internal states correlated to structures of the cognitive system that evaluate it. There is no space here in the present paper to deepen this discussion, but it is opportune to say that our own experiments showed the need for an emotional system that serves as the motor for the agent's actions. According to [Rocha Costa & Graçaliz 2003], internal motivation is the basis to real autonomy of agents. This result is according to recent studies of cognitive research where it is claimed the mutual dependence between knowledge and affectivity in the intelligent behavior. Body and emotional system are fundamental elements to guarantee the autonomy of our constructivist agent, thus all intentionality is an emergent result of its own mechanisms. This claim have support in [Bercht 2001], [Sloman 1999], [LeDoux 1996] and [Damásio 1994].

The triple {Context, Action, Expectation} is composed of vectors of elements that may undertake three values: true, false or don't care (represented, respectively, by '1', '0' and '#'). A true value indicates the presence of an element, a false value indicates absence and the third value is to make a generalization of the element in the schema, don't care indicates no matter if it is present or not. For example, a schema that has the context vector = (0,1,#) is able to assimilate situations (0,1,0) e (0,1,1).

There is compatibility between a schema and a certain situation when the context vector of the schema has all true or false elements equal to those of the agent's perception vector. Note that compatibility does not compare the 'don't care' elements.

After an activation, the schema is re-evaluated. Its evaluation measures the prediction ability of the system. A schema has a good capacity of predicting if its expectations are fully corresponded after the execution of its action, and can then be called 'adequate schema'. Evaluation happens by the analysis of compatibility between the expectation vector of the schema applied at the previous instant, and the agent's perception vector. As this fitness is historic, and accumulates during the several applications of the schema, it is expected that it tends to a value that describes its real prediction ability.

4. CONSTRUCTIVIST AGENT: ALGORITHMS

Schemas are always re-evaluated after each activation. The new fitness of a schema should be equal to some ponderation between its previous fitness and the expectation compatibility of the schema with the current situation. For instance, as indicated:

```
RE-EVALUATION Method
  S := ActivatedScheme;
  S.NewFitness := (S.OldFitness + Compat(Scheme.Expect))/2;
```

The learning process of our constructivist agent happens through the refinement of the set of schemas, this means making it more adapted to its environment. This learning happens through four basic methods, named: Guided Creation, Expectation Generalization, Context Generalization and Differentiation.

When the agent faces a situation where it does not have compatible schemas (or if those that it has are, in some way, rejected by the emotional system) then it activates the Guided Creation method, with the aim of widening its repertoire of schemas and dealing specifically with the situation it is at that very moment. The method works as it follows:

```
GUIDED CREATION Method
  if Agent.ActivatedScheme = null then
    S := Tscheme.Create;
    S.Context  := Agent.Perception;
```

```
S.Action := NewActionForContext;
Agent.AddScheme(S);
Agent.Execute(S);
S.Expect := Agent.Perception;
S.Fitness := 0.5;
```

The whole new schema created by the Guided Creation method has the vector of context and Expectation totally specific, created after the situation experienced and after the result directly observed after the action. The cognitive mechanism of the agent has methods to find the most generic schemas as possible after these initial specific schemas, adjusting the expectations and integrating analogue contexts. In more complex environments, the number of sensations the agent realizes is huge, and, in general, from these, only a few of them are relevant in the context of a schema. In the same way, an action the agent performs will generally modify only a few aspects of the environment.

Expectation Generalization modifies an existent schema in order to make it more adequate, that is, more precise in its expectations. It works as a kind of accommodation, once the schema adjusts itself according to its last application. The method simply compares the activated schema's expectation and the agent's perception after the application of the schema and changes the elements that are not compatible to each other for the indefinite value '#'. As the Guided Creation method always creates the schema with expectations totally determined and equal to the result of its first application, the walk the schema performs is a reduction of expectations, up to the point it reaches a state where only those elements that really represent the regular results of the action carried out in that context remain.

```
EXPECTATION GENERALIZATION Method
  S := Agent.ActivatedScheme;
  For i := 0 to Agent.Perception.Size-1 do
    if S.Expect[i] <> Agent.Perception[i] then
      S.Expect[i] := #;
```

When the agent finds two similar schemas to approach different contexts, the Integration algorithms come into action. These two schemas only trigger the procedure if they have high fitness (being considered reliable) and if they have the same action and expectation (action and expectation vectors must be completely compatible). As all schemas are born with their context very well specified, the mechanism needs to find the state with the higher level of generalization as possible, without loosing adequacy. The algorithm operates as it follows:

```
INTEGRATION Method
  if (S1.Action = S2.Action) and (S1.Expect = S2.Expect)
then
    if (S1.Fitness = 1) and (S2.Fitness = 1) then
      S := TScheme.Create;
      S.Action := S1.Action;
      S.Expect := S1.Expect;
      For i := 0 to S.Context.Size-1 do
        if S1.Context[i] <> S2.Context[i] then
          S.Context[i] := S1.Context[i]
        else
          S.Context[i] := #;
      Agent.SubstituteSchemes(S1, S2, S);
```

Differentiation becomes a necessary mechanism because Generalization may make inapropriate generalizations, occasionally. If a generalized schema does not work well, it creates a new schema. It acts as it follows:

```
DIFFERENTIATION Method
  S := Agent.ActivatedScheme;
  if (S.Fitness = 1) and (S.InstantFitness < 1) then
    NewS := TScheme.Create;
    NewS.Context := Agent.LastPerception;
    NewS.Action := S.Action;
    NewS.Expect := Agent.ActualPerception;
```

Thus, the algorithm that chooses the schema that will be activated at each turn in the agent should be prepared to choose always the compatible schema that has the most specific context.

5. MODEL SIGNIFICATION

The problem can be stated in the following form: the environment has certain regularities that the agent may check during interaction. Considering that these regularities are stable, there will be, for an agent which is in this environment, a set of valid schemas that the agent should be able to build.

If we consider context, action and expectation as a triple search space, then the problem solving means, for the constructivist agent, to find the set of valid schemas in this search space. Thus, after some time, each schema of the set of schemas that the agent has built should represent a valid correlation between a context, an action and an expectation in the given environment. The set of built schemas defines adaptation of agent to its environment.

The agent model we have mentioned here has a reactive profile: although it learns with experience, it does not have temporality (it does not take into consideration past through any kind of historic memory, and does not plan the future), thus limiting itself to understand only instant relations; it does not count on a symbolic capacity nor abstraction, that is, it is not able to formulate concepts that go beyond the perception level, sensorimotor.

The whole new schema is generated through a process of Guided Creation. This method creates a schema absolutely specific, because its Context vector is only able to assimilate this unique situation, and its expectation vector expects to find always the same exact result after the action application. The cognitive mechanism finds an adequate generalization of this schema through the Expectation Generalization, Integration and Differentiation methods.

6. PRELIMINARY RESULTS

We have made an experience of implementation of the Constructivist Agent, inserted in a simple virtual environment. In the simulation, the agent should learn how to move in a plane without colliding with obstacles. The plane is a bidimensional grid. The cells of grid may contain an empty space or a wall. At each time, the agent can do one of three possible actions: give a step ahead, turn to left or turn to right. The agent walks freely across empty spaces, but collides when tries to step against the wall. Figure 3 shows the simulation environment, where dark cells are walls, clear cells are empty spaces, and the blue cell is the agent, seeing one cell ahead.

Initially the agent does not have any schema and it also did not distinguish the obstacles from the free ways. The agent had only one external perception: the vision of what was in front of it, and the sensation of pain when there happened a collision. The agent's body has four properties: *pain, fatigue, exhaustion* and *pleasure.* All of these properties have corresponding internal perceptions. Pain ocurrs when the agent collides with a wall, and lasts just one instant. Agent's corporal metabolism eliminates the sensation of pain at the next instant. When the agent repeats much times the action of walk, then fatigue is activated, and if the walk continues, sensation of exhaustion is finnaly activated. These sensations disappear after a few instants when the action of walk ends. Pleasure occurs when the agent walks, during only one instant too.

Agent's emotional system implements three emotional triggers: pain and exhaustion define negative emotional values, whereas pleasure defines a positive value. At the end of a period of interaction with the environment, we hope the agent had been able of building a set of schemas that prevented it of making the moves that leads to pain in each visual context, and prevent it of exhaustion when fatigue appears. In addition, the agent prefers walking in other situations, because it feels pleasure.

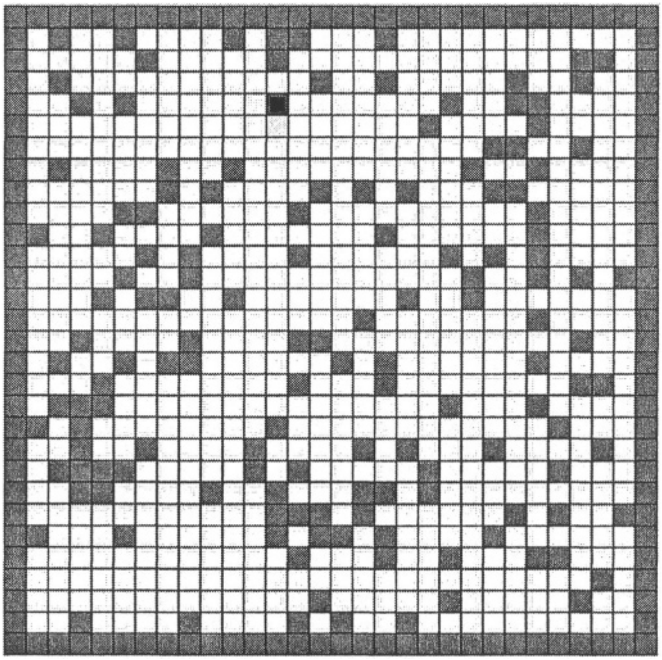

Figure 1. Simulation Environment

After some experiments, we may consider our results were successful. The agent, after some time in interaction with the environment, learns about the consequences of its actions in different situations. This knowledge is used by the agent to avoid emotionally negative situations, and to pursue emotionally positive situations. In Table 1 we show the main schemas built by the agent in the simulation.

***Table 1.* Main schemas built by the agent during the simulation**

Context	Action	Expectation
* vision = wall	* walk ahead	* pain = true
		* vision = wall
* (any context)	* turn left	* pain = false
	* turn right	
* fatigue = true	* walk ahead	* exhaustion = true
* fatigue = false	* walk ahead	* pleasure = true
* vision = empty space		

7. CONCLUSIONS AND NEXT STEPS

The constructivist models in AI are interesting as they make possible to develop architectures of intelligent agents that adapt themselves to the environment without the programmer's intervention in the construction of its cognitive structures. The agent, although programmed to accomplish certain tasks, is free to build its knowledge in the interaction with the environment, and thus, finding alone the solutions for problems that may arise. This autonomy of cognitive construction is associated to a motivational autonomy, given by the corporal and emotional sensations.

The model presented hitherto is an alternative that follows these guidelines. The achieved experimental results show that the constructivist agent is able to learn regularities of its environment. After some time, the agent behaviour becomes coherent with its emotional values because a set of adapted schemas was constructed through the continual interaction with the environment.

However, it is limited to a too reactive agent (without memory, nor planning, nor conceptual abstraction). We hope, as a continuation of this work, to point a way for the implementation of a more cognitive constructivist agent that has temporality and conceptualization. Everything indicates that the next step means to make possible that the agent builds multiple sets of schemas related to each other, making the meaning of perception and acting vary in a way that these also make reference to their own cognitive activity.

Constructivist Psychology is in accordance with this systemic perspective of the mind. Furthermore, the idea of a modular mind operation (where each module works in a specific domain, but in constant interaction with other modules) is postulated by several researchers in neuropsychology [Rapp 2001], [Luria 1979], [Pinker 1997], [Damasio 1996]. Also in AI, this Idea can be found, somehow, in [Minsky 1985].

In this way, the agent architecture will be extended to integrate the following capabilities:

- temporal perception: the agent must be able to perceive time distributed contexts, in addition to the capability of identifying regularities in the instantaneous perception. Thus, the agent will observe sequenciated events; this capability is requirement to emergence of planning and factual memories.
- abstract concept formation: the agent must be able to overpass the sensorial perception constructing abstract concepts; forming this new knowledge elements, the agent aquires new forms to structure the comprehension of its own reality in more complex levels.

- multiple knowledge systems: the agent must be able to construct subsets of perceptions or concepts, which represent specific knowledge domains; a domain is a dynamic knowledge system. In the interaction, these domains represent proper significances mediating the specific agent adaptation.

- operatory transformation: the agent must be able to create operations that change among different systems or schemes, preventing that knowledge remains as a database of cases, and enabling the agent to understand further phenomenon regularities, systematic transformation regularities in the world.

- non-modal logics: the agent cannot have only deductive inference. A hypothesis inference mechanism is required, as well as uncertainty and inconsistency treatment.

REFERENCES

Becker, Fernando. Educação e Construção do Conhecimento. Porto Alegre: ArtMed, 2001.

Belpaeme, T., Steels, L., & van Looveren, J. The construction and acquisition of visual categories. In Birk, A. and Demiris, J., editors, Learning Robots, Proceedings of the EWLR-6, Lecture Notes on Artificial Intelligence 1545. Springer. 1998.

Bercht, Magda. Em direção a agentes pedagógicos com dimensões afetivas. Porto Alegre: UFRGS, 2001. (Tese de Doutorado)

Birk, Andreas & Paul, Wolfgang. Schemas and Genetic Programming. Ritter, Cruse, Dean (Eds.), Prerational Intelligence: Adaptive Behavior and Intelligent Systems without Symbols and Logic, Volume II, Studies in Cognitive Systems 36, Kluwer, 2000.

Boden, Margaret. Piaget. Glasgow: Fontana Paperbacks, 1979.

Coelho, Helder. Sonho e Razão. 2.a. ed. Lisboa: Editora Relógio d'Água, 1999.

Damasio, António. Descartes' Error: Emotion, Reason and the Human Brain. New York: Avon Books, 1994.

Davidsson, Paul. On the concept of concept in the context of autonomous agents. II World Conference on the Fundamentals of Artificial Intelligence. p85-96, 1995.

De Jong, E. D. Autonomous Concept Formation. In: Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence IJCAI, 1999.

Drescher, Gary. Mide-Up Minds: A Construtivist Approach to Artificial Intelligence. MIT Press, 1991.

Holland, John. Genetic Algorithms. Scientific American, 1992.

Inhelder B. & Cellerier G. Le cheminement des découvertes de l'enfant. Neuchâtel: Delachaux et Niestlé, 1992.

Kohonen, Teuvo. Self-Organization and Associative Memory. Berlin: Springer-Verlag, 1989.

LeDoux, J.E. The Emotional Brain. New York: Simon and Schuster, 1996.

Luria, Aleksandr. The Making of Mind. Harvard University Press, 1979.

Minsky, Marvin. The Society of Mind. New York: Simon & Schuster, 1985.

Montangero, J., & Maurice-Naville, D. Piaget or the advance of knowledge. New York: Lawrence. Erlbaum Associates, 1997.

Muñoz, Mauro H. S. Proposta de Modelo Sensório Cognitivo inspirado na Teoria de Jean Piaget. Porto Alegre: PGCC/UFRGS, 1999. (Dissertação de Mestrado)

Piaget, Jean. Play, Dreams and Imitation in Childhood. London: Heinemann, 1951.

Piaget, Jean. The Origins of Intelligence in the Child. London: Routledge & Kegan Paul, 1953.

Piaget, Jean. Construction of Reality in the Child. London: Routledge & Kegan Paul, 1957.

Piaget, Jean. A Epistemologia Genética; Sabedoria e Ilusões da filosofia; Problemas de Psicologia Genética; Vida e Obra. (Os Pensadores). São Paulo: Abril Cultural, 1978.

Steven Pinker. How the Mind Works. W. W. Norton, New York, NY, 1997.

Rapp, B. (Ed.). The Handbook of Cognitive Neuropsychology. Hove, UK: Psychology Press, 2001.

Rocha Costa, Antônio Carlos da, & Dimuro, Graçaliz Pereira. Needs and Functional Foundation of Agent Autonomy. 2003. (http://gmc.ucpel.tche.br/imqd/artigos/needs-03-04-06.pdf, 21/1/04)

Rocha Costa, Antônio Carlos da. Inteligência Artificial Construtivista: princípios gerais e perspectivas de cooperação com a informática na educação. Porto Alegre: Instituto de Informática UFGRS,1995.

Sloman, Aaron. Review of Affective Computing. AI Magazine 20 (1): 127-133, 1999.

Wazlawick, Raul. Um Modelo Operatório Para a Construção do Conhecimento. Florianópolis, PPGEP/UFSC, 1993. (Tese de Doutoramento)