Learning in technical systems: a Sign based approach

Jaime Gómez Ricardo Sanz and Ramón Galán

Abstract— **In this paper the following assumptions are conveniently explored**

i) there is no cognition without categorization

ii) concept formation is sign a mediated process which is extended and developed in a formal and operational fashion.

iii) learning is involved with forward and inverse cognition.

A new framework for knowledge representation in cognitive agents is proposed, whereby the formal and operational issues of representation are faced through a systemic conception of *sign*.

As a result of the systemic and integrative approach taken, the basis of a cognitive theory for concept learning is set up.

I. Introduction

The main function of cognition is categorization. By categorization we mean the action of grouping the perceived items into the class. [11][12][22][18]

Categorization and cognition are here acknowledged as equivalent terms, any cognitive system has to be able to categorize. Within this rubric, a cognitive system can be biological, from superior mammals to simple bacterias, or artificial like cognitive robots. They are all agents able to categorize, ergo cognitive agents.

This claim is untenable from the doctrine that assumes that it is necessary to have a neural system in order to have cognitive abilities. Under this assumption, only animals with a sufficiently developed brain posses concepts and may be adaptable agents in a changeable environment[3].

The positioning of this paper is radically different from that view. The identification of cognition and all its properties like memory, learning, representation etc. in the human brain, is a relic of dualism dressed with topics of human chauvinism.

A technical system, for example a cognitive robot, is not a system deprived of sensors and confined in a unique context, rather it is an agent with the ability to cope succesfully with a dynamic environment.

The approach followed in this work is drastically opposed with representation as the encoding of that which is represented. Such strategy assumes that knowledge is synonymous to representation, consequently the knowledge about the world would consist of the sum of sentences empirically verifiable. Assuming that, Wittgenstein would be right in his famed assertion that "the limits of my language mean the limits of my world"[25].

This work has been supported by ICEA (Integrating Cognition, Emotion and Autonomy) funded by IST Cognitive Systems Unit.

Jaime Gómez and Ricardo Sanz are with ASLab, Universidad Politécnica de Madrid, Jose Gutiérrez Abascal 2, Madrid 28006, Spain. jagomez@etsii.upm.es

Ramón Galán is with Intelliget Control Group, Universidad Politécnica de Madrid, Jose Gutiérrez Abascal 2, Madrid 28006, Spain.

We are closer to the thesis defended by Bickhard that conceives that knowledge is the ability to successfully interact with the environment[2].

Although we find Bickhard's interactive representation of knowledge inspiring, we still are of the opinion that it is too vague and hardly applicable to technical systems.

The thesis we propose is radically different from those views. The cognitive agent is conceived as a "sign hunter", counting on a repertory of sensors that enables it to acquire loose input patterns from the environment and embed them into signs.

II. The problem of representation

For the sake of clarity, we will begin with some definitions.

Etymologically, to re-present is to make something present, bring before the mind[15] or to capture the structure of that which is re-presented. It is worth noting that the regular use of the word is wrought with a pictorial idea of representation as a roughly accurate copy of reality.

Heylighen defines representation as *"a formal structure relating observable phenomena; as such, it provides a mechanism for deriving anticipated observations (output) from actual observations (input)"*[13]. Of course, the external phenomena can be a physical or an abstract thing. For example, a sculpture in Madame Toussaud's wax museum or a plastic toy rhinoceros, are representations of physical things. Similarly, the word "love", written on ink or pixels, is a representation of an abstract thing.

Knowledge representation is a set of ontological commitments that declares the category of existing things and the process undergone by things in a given domain. The ontological commitments are formally expressed through a language or collection of propositions logically related, which defines a model of the domain.

Traditionally, the problem of representation has been conceived as a diadic structure. On the one hand the thing to be represented and on the other the symbol that represents the thing.

The classical theory is based on the aristotelian tradition that establish that an item i fall under a concept C, $i \in C$, iff i satisfy a set of necessary and sufficient properties p.

Formally,
$\nexists j | p(i)p(j), p(j) \bigcap C \neq \emptyset$.

Such a conception assumes the existence of a morphism between the external world and the agent's world model.

Alternatively, the interactive representation model (Bickhard), is characterized by the emphasizes on features such as action and expectation, disregarding the formal structures that tend to attain an isomorphic representation of reality.

In the classical theory of representation, knowledge about the world is a static picture composed of atomic entities that can be formally described with propositions.

Operationally this is how it works; the internal objects of the agent are represented by symbols that commit correlations with the external objects. When the correspondence between internal (symbols) and external objects (referents) is an isomorphism[1], a perfect correlation between both set of objects is achieved. In this case, the agent has an ideal representation of the environment.

But this representation structure is unsuitable if we want to build an artificial agent that copes successfully with the external world. This is because the world is not static but inherently dynamic and representation has to take into account the fundamental categories of change and time. Secondly, an isomorphism with the environment can only be attained duplicating the environment, only in this case there is not singular points, put simply for every pair of things in the external world, there are their correlates in the agent's model. But this is a totally intractable strategy for building cognitive agents.

In order to cope with an inherently dynamic and mutable world, the cognizer needs of adaptative meta representational capabilities. The ontology of concepts and their relations in order to be operational, can not be invariant but malleable.

III. Building a new Theory of Cognition based on Signs

The theory of representation shown above, is dependent upon insufficiencies as a consequence of the diadic structure of representation, composed by the symbol and the concept the symbol refers to, expressed by the tupla:

$< symbol, concept >$.

The trouble with such an approach, even if it can succeed in the formalization issue is that it lacks an operational asset. The representation of relations between the internal model of the cognitive agent and the environment is dynamic, situated, evolving and strongly interconnected with the cognizer. Thus, it can not be defined following the formalization method used in classical theory, by which the membership of an item to a category is unambiguously established by a a set of axioms and deductive rules.

On the other hand, interactive theory, which claims that concepts are acquired dynamically, embodied and inherently linked with senso motor components, supplies better operational capabilities but the formalization is hardly attainable.

The cognitive theory proposed here is inspired by Peirce's triadic structures formulation designated as a sign. In our theory, a sign is a system composed by the next three elements: symbol, concept and referent. Formally, a sign S is the ternary tuple

$S =< symbol, concept, referent >$

The sign's components are described as follows:

- Symbols are the most basic component of a sign, their function is to share concepts in a community of agents by means of a consensual language symbols denoting signs.
- Concepts are mental models nomologically related. Concepts can be viewed like the intension of a sign.
- Referents are the objects and their properties that the cognitive agent is capable to identify based on both the agent's sensorial properties and the ontology of concepts possessed by the agent. Referents are the extension of a sign.

[1]Let α, β realizations of the same type of structure.An isomorphism $\Phi : \alpha \rightarrow \beta$ is a structure preserving mapping 1:1 of α onto β

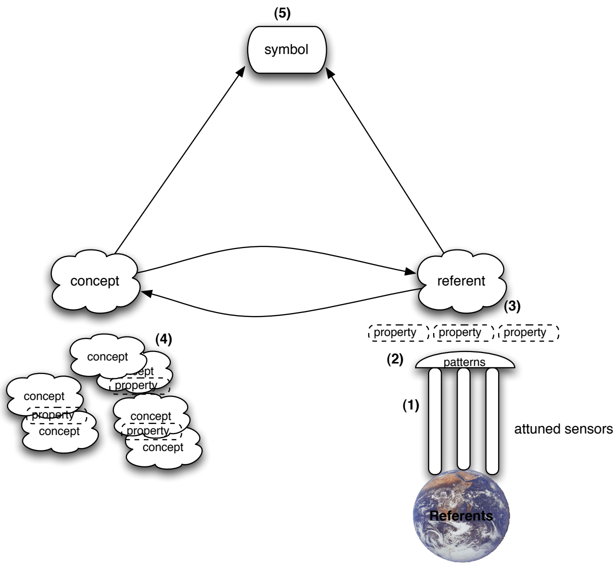

Fig. 1. Sign structured composed by three components: symbol, concept and external referent

The components of a sign are not basic primitives but they are systems themselves as well. The simplest component of a sign is the symbol which can function *in the head of the cognitive system* as an icon, depicting one external object, or as an index that enable the agent to refer a concept and share concepts within a community.

According to this, signs are not symbols. Rather, a symbol is a sign's component, which is amodal and uni-dimensional whereas a sign is multimodal and N-dimensional.

Hence, we now reformulate Peirce's definition of a sign *"signs are something which stands to somebody for something in some respect or capacity"* by "signs are systemic structures which stands to somebody (who posses the concept), thorough the sign's component symbol, for something(referent) in some respect or capacity(concept)".

In [8] we propose, though incomplete, a naturalized theory of concepts which are seen as embodied mental representations. Three different levels of cognition are explored: linguistic, sub-linguistic and neural.

In [9] some important steps are given towards an explanation of how the neural structures can ascribe meaning in order to get the causal/computational isomorphism between neural structures and their correlated concepts.

If we assume the hypothesis that the nature of concepts is inherently multilevel and systemic, then language per se, is not a necessary piece in a theory of cognition that explains how we learn and categorize new concepts.

The search for a language of thought is a path that can only give bias answers to questions such as how the cognitive agent cognizes and what do we need to build artificial cognitive machines. We, like Eliasmith hold that "..starting with language or even focusing on language, when constructing a theory of content is a dubious tactic"[5].

A. *The process of categorizing for a physical object*

Now we attempt to to explain the process of categorizing for a physical object, as is depicted in Figure 1.

Let start upon the first step in the process of cognition, perception. The referents of the external world are grasped by the agent. This is because the world is populated by material things which undergo processes which emit energy to be captured and encoded by the sensors, as is shown in (Figure 1(1)). The sensory stimuli captured by the agent are objective and quantifiable. The properties of the perceived object can be measured, of course, the agent has perceptual limitations about what can and cannot perceived, based on its sensors and the way they are attuned.

The patterns are instantiations of concept's properties for certain kinds of perceptions (figure 1 (2)) that try to achieve the matching with the encoded information of the sensor's channels (Figure 1 (3)). When this computation succeeds, the referent is incorporated to the concept ontology. In other words, the salient features or properties of the referent are identified and related to the agent's ontology of concepts.

The conceptual component of a sign is depicted in (Figure 1 (4)). In actual fact, it is an ontology of concepts which represent things or processes with common properties. According to this, the ontology of concepts is nomologically related by the relationship among the properties of the concepts.

Due to the lawfulness of the concepts relations, learning is possible; if the brain lacked this, that is to say, the properties that belong to the concepts the perceived item from the external world could never be classified. There it would be an agent with deficient cognition and scattered options to survive in a world ruled by laws.

Alternativelly, if the agent, as is the case in humans, has a language or some other sign-denotative system of symbols, the relation between the external referent and the ontology of concepts can be by- passed by a symbol. The symbol (Figure 1 (5)) serves as a vehicle to share concepts within a community of agents. For example, a kid can learn to grasp something before the mother gives him the symbol "grasp" as the label that identifies the sign grasp that represents an action in a community of common language speakers. It goes without saying that the kid could have the sign grasp without the symbol "grasp", he can have the concept in his head and can discern when someone is grasping something but be unable to utter "grasp" or react when hears such word.

These observations convey to us a new definition of representation as *the process of construction and relationship of signs within a system of signs*.

Accordingly, learning is to relate signs to system of signs.

Signs are therefore interpretations of patterns that tend to capture features of both the environment and the own cognitive agent. The cognizer through categorization can make predictions about objects and events. The evolutive advantage of anticipation has been extensively discussed by Rosen[20] and Deacon[4] suggests a parallel between conceptualization and Evolution Theory, "concepts evolve in social environments in much the same way that organisms in natural *Ulmwelt*".

B. *A solution for the symbol-grounding problem*

The so called symbol-grounding problem [10], arises from the supposition that computers are disembodied machines, processing symbols arranged as is established by a set of syntactic rules, and consequently incapable to deal with conceptual content and semantics.

The symbols, according to Harnad, are positively disconnected to the things they denote; this is because the symbolic computation is totally decontextualized, it does not depend on the medium in which it is implemented.

But the problem vanishes if we recognize the sign-mediated nature of cognition. The cognitive agent, is able to interpret the environment only when satisfies next requirements:

i It is sensorily attuned to the properties of the perceived thing (Figure 1 (1)) and

ii It has models reified in the patterns used to match the salient features of the perceived things (figure 1(2))

The system of signs, guides perception, categorization and the motor action of the agent. To summarize, when matching pattern succeeds (Figure 1 (3)), a sign is created and an external phenomena and a portion of the internal model of the agent are now linked.

IV. Learning concepts in artificial systems

A. *Gardenfords' conceptual spaces*

Gardenfords[6] proposes to use conceptual spaces as an alternative mode of representation to symbolic and associationist accounts. Gardenfords' model is based on two main principles:

- *Economy principle*: the items with similar characteristics fall under common categories. Operationally, this means that the items are stored in a common memory space which defines a range of properties and their admissible values.
- *Non-monotonicity effects in reasoning*: it seems that the similarity between two items that belong to a concept cannot always be expressed in terms of necessary and sufficient conditions, but is based on prototype structures. Thus, the membership of an item to a concept is a question of degree; the closer the item is to the gravity center of the prototype structure (the dots in Figure 2), the better representative of the category it is.

Voronoi's diagram (Figure 2) defines a geometric space that represents the prototype space of a concept. Following this mathematical algorithm, the clusters are constructed dynamically as long as new items are cognized. The explanatory power of this approach is undeniable, and its easy formalization and implementation makes of Voronoi's tessellations use in conceptual learning, a good candidate for a concept learning model.

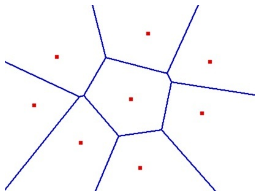

Fig. 2. A Voronoi diagram is decomposition of a planar metric space determined by distances to a discrete set of points, in our case representing items to be categorized. The simplest case is a plane, only two dimensions with a number of regions that configure the different concepts where the items can fall under.

In the practical case this model of learning can hardly operational. If the cognitive agent is able to rely on multiple and more accurate sensors, the number of properties increases, and as every property, following Gardenford's model is represented by an axis; if we have the very plausible number of four properties, the Voronoi's diagram becomes practically illegible, let alone if we continue increasing the number of axis.

Recall that the domain is the set of dimensions that in turn are the observable properties represented in axis. Of course, the more properties are discernible by the agent the more accurately the item can be categorized.

But the trouble here is that counting on more properties does not necessarily help to categorize items in a more accurate way, rather the proliferation of dimensions (properties) makes the problem intractable.

B. Learning concepts in a theory of cognition sign based

In the previous section we presented a general purpose theory of cognition. The theory emphasizes the triadic structure of sign as the minimum structure of cognition and provides a theoretical framework where models of particular cognitive abilities like learning can be implemented and tested in artificial machines.

In our model, for the sake of simplicity, we set aside properties that are not directly generated by the sensors like second order properties (qualia). That being said, is important to keep in mind the fact that there are properties of a more complex nature, we call them phenomenal properties similar to Galileo's second order properties. Phenomenal properties only exist in the "head" of the cognitive agent, that is in absence of a cognizer there is no such things as heat, redness or fear.

On the contrary, the mass of a body or its velocity exists by its own and is almost independent of the agent observing it (totally independent if we exclude atomic interactions).

Figure 3) shows a general diagram of the control processes involved in learning. Forward cognition consists of representing in some physical support (biological or not), the properties of the external referent.

Thus, in Figure 3, $P = p_1, p_2, p_3$ are the properties of the external object perceived by a technical system which are represented in its physical substratum as $P = p'_1, p'_2, p'_3$

The properties captured by the agent determine the content or semantic interpretation of the referent. The process of learning starts when the properties of the referent are perceived and measured by the cognitive agent. Why are these and no other properties of the referent grasped by the agent, or how the perceptive process is achieved by the agent is a theme for another paper, we are focusing here on learning and categorization and not in perceptual issues.

Once the agent recognizes and measures the significant properties of the referent, Forward Cognition box in Figure 3, then the agent has a representation of the referent, given by its measured properties $P = (p'_1, p'_2, p'_3)$.

The next step is to categorize the item perceived and measured by its properties, Inverse Cognition box in Figure 3. That is to say, the items fall under concepts or categories which express the space state that defines the lawfulness of the relation between the properties.

Formally this can be expressed as follows:

Let P, the properties measured for an item i which belongs to the category or concept ζ.

$P = (p_1, p_2, ...p_n) \in \zeta$

and let φ, the Forward Cognition function implemented in the technical system, then

$\varphi(p_1, p_2, ...p_n)$=$(p'_1, p'_2, p'_3)$

Consequently, the categorization process is formally expressed by the inverse function φ^{-1}.

$\varphi^{-1}(p'_1, p'_2, p'_3)$=$(p_1, p_2, ...p_n) \in \zeta$

Forward Cognition is assimilated with perception and Inverse Cognition with learning. We can determine, mostly a priori, what and how a technical system with sensors perceive the external objects. That is, the function φ that represents the forward cognition of the system, can be calculated through empirical studies, namely $\triangle = \varphi(P) - P$ measures how accurate is φ.

Alternatively φ^{-1} is the implementation of learning[2];

$\varphi^{-1}(p'_1, p'_2, ...p'_n)$=$(p''_1, p''_2, ...p''_n) \in \zeta$ iff $\triangle = (P)'' - (P)$ is minimal. Put in another way, the technical system learns that an item belongs to a category ζ when the error in the measured properties $\triangle$ is smaller than for any other category.

For the principle of economy, the agent places each item into a category, otherwise every single represented item, would be different and unrelated with the rest of items and the agent would perceive the lawfulness of the external world as a chaos.

C. An example of learning in a technical system

To this point, we can say more about how a technical system categorizes items. Recall that our theory is a general

[2]Note that the obtention of φ^{-1} is only pertinent in the ideal case of isomorphism between the mental model of the agent and the perceived object in the world

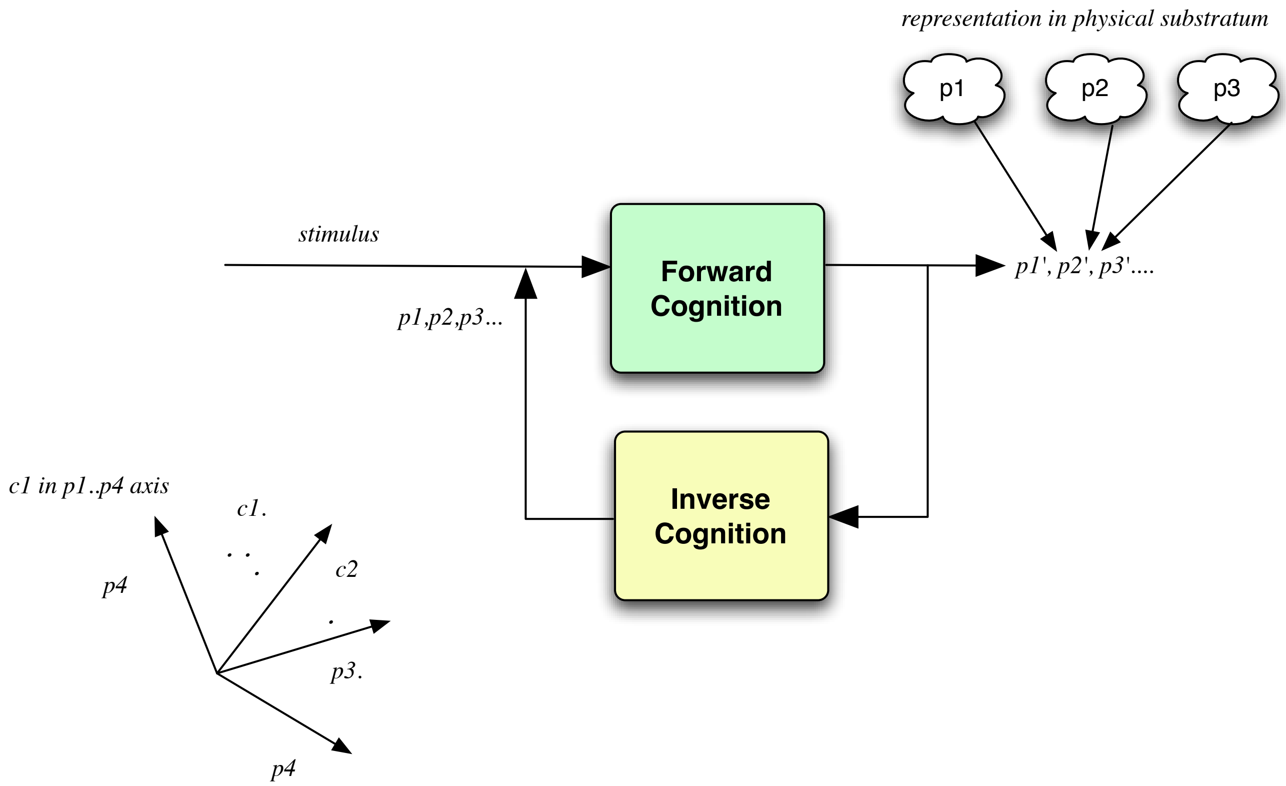

Fig. 3. Learning is involved with forward and inverse cognition, the properties of the perceived thing are represented in a physical substratum.The more accurate is the representation of the perceive properties the less is the categorization error

and systemic framework for cognition and therefore is appropriate for building a model of concept learning referred to physical objects.

For illustrative purposes, let us consider an scenario with a cognitive robot and an object moving in circles around it. The robot count on the necessary sensors to measure the next properties of the object:

mass*(m)*, angular moment *(L)*, velocity*(v)*, acceleration*(a)* and distance*(d)*.

Thus, the perceived objects by the robot are represented in a four dimension domain which compose the system of reference. One dimension or axe in the system of reference, represents the measured values of a property.

In Figure 3, the concepts c1 and c2 are depicted, based on the quantitative and objective measurements of the relevant properties.

Suppose now that the robot is confronted with a first object,its sensors measure the relevant properties in three different instants with the next results:

For t=0:

$P(t_0)$=(m=1kg,L=1kg$m^2 s^{-1}$,v=1ms^{-1},a=0ms^{-2},d=1mm).

$\varphi^{-1}(P(t_0)) \in \zeta$, where ζ is the $item_1$ that follows a rotational trajectory ($L \neq 0$) with acceleration null. Therefore the first item is categorized in the category ζ which is equals to the measured item, $\zeta = item_1$

For t=1:

$P(t_1)$=(m=1kg,L=1kg$m^2 s^{-1}$,v=2ms^{-1},a=0ms^{-2},d=1mm).

$\varphi^{-1}(P(t_1)) \in \zeta$, the item still has a rotational trajectory with acceleration null. Therefore for the principle of economy it is placed the concept ζ

For t=2:

$P(t_2)$=(m=2kg,L=0k$m^2 s^{-1}$,v=2s^{-1},a=1ms^{-2},d=1 mm).

$\varphi^{-1}(P(t_1) \in \xi$, a new category is created, the item has not rotational trajectory ($L = 0$) and is undergoing a force ($a \neq 0$).

The system of signs constructed by the robot is as follows.

$S =< symbol, concept, referent >$ then,

$S_0 =< item_0, \varphi^{-1}(P(t_0)), \varphi(P(t_0)) >$

$S_1 =< item_0, \varphi^{-1}(P(t_1)), \varphi(P(t_1)) >$

$S_2 =< item_2, \varphi^{-1}(P(t_2)), \varphi(P(t_2)) >$

Then, S_0 and S_1 are two measurements, in two different instants of time, of the same object, $item_0$, which belongs to the class ζ, that groups objects with rotational trajectory and null acceleration.

The sign S_2 has as concept component a new category ξ. Indeed, the perceived object namely "$item_2$", has a different mass value, and its trajectory is accelerated and not rotational.

Note that the attribution of class in this example has not been defined here because we are presenting a general purpose theory open to any the applicative use.

Thus, for example a robot that has to deal with moving objects will categorize objects based in properties as velocity and one robot evolving in a static environment will develop categories from properties as mass or distance.

The advantage of the systemic and semiotic approach presented here is that the technical agent can builds a knowledge base that reckons the inherent dynamic and malleability of the real world.

V. CONCLUDING REMARKS

Accordingly, while our theory is still in its embryogenic stage we can not give a formal account of the learning process. However at the present stage we can provide the next hints that will be tested in following works.

- every property is a property of a thing, there are no properties by themselves, things undergo properties.
- a representational system is not necessarily a neural system. The symbol grounding problem is a fake problem or better said, a problem motivated for the classic diadic conception of representation.
- learning processes require both forward and inverse cognitive processes. In a technical system, the former, formally φ, measures the the relevant properties for the perceived object, and the last, φ^{-1}, through a transformational procedure, group the item, represented by its measured properties, into a class of items.
- the transformational procedure is a function φ that maps physical objects into the domain of properties, the learning process is complete only when φ^{-1} is determined; that is, the function maps the representation of the properties (in the case of mammals the firing of the neurons) into the class where the thing belongs. Note that φ^{-1} is the formalization of the categorization process or learning.

It is worth noting that concept learning and categorization are a difficult and fundamental problem in cognitive science.It is not a problem that anyone should expect to solve based only on theoretical assets about cognition.

Science must proceed testing the theoretical framework against facts, empirically testable.

Eventually this theoretical effort will bear fruit a new way of looking at things that changes the state of the field in learning and cognition in artificial systems.

REFERENCES

[1] Adamek, J. Herrlich H. Strecker G.E, (1990) Abstract and Concrete Categories, John Wiley and Sons.

[2] Bickhard,M.H. (1999) Interaction and Representation, Theory s Psychology, Vol. 9, No. 4, pp. 435-458

[3] Bunge, M. (2008) Blushing and the Philosophy of Mind, Journal of Physiology-Paris

[4] Deacon, T. (1997). The symbolic species: The co-evolution of language and the brain, Norton.

[5] Eliasmith, C. (2007) How to build a brain: From function to implementation, Synthese.

[6] Gardenfors, P. (2004) Conceptual Spaces: The Geometry of Thought, MIT Press

[7] Goguen,J. (2004) Ontotheology, Ontology, and Society, In Formal Ontology in Information Systems, edited by Achille Varzi and Laure Vieu, IOS Press, pp. 95-103

[8] Gomez J. et al. (2007) Naturalized epistemology for autonomous systems, Kazimierz Naturalised Epistemology Workshop

[9] Gomez J. et al. (2008) Cognitive Ontologies: Mapping structure and function of the brain from a systemic view, accepted at AAAI 2008 Fall Symposium on Biologically Inspired Cognitive Architectures

[10] Harnad, S. (1990). The symbol grounding problem, Physica D 42, 335-346.

[11] Harnad, S. (2003) Categorical Perception. Encyclopedia of Cognitive Science. Nature Publishing Group. Macmillan

[12] Harnad, S. (2005) To Cognize is to Categorize: Cognition is Categorization in Handbook of categorization and cognitive science,

[13] Heylighen, F. (1990) Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science, Communication Cognition, pp. 200.

[14] Neuman, Y., Nave, O. (2007) A mathematical theory of sign-mediated concept formation, Applied Mathematics and Computation.

[15] The Concise Oxford Dictionary of English Etymology (1996), Oxford University Press.

[16] Patterson, S. (1998), Competence and the Classical Cascade: A Reply to Franks, The British Journal for the Philosophy of Science.

[17] Peirce C.S. Collected Papers of C. S. Peirce, Harvard University Press, Cambridge, MA, 1931-1958.

[18] Rey G. (2005) Empty Concepts and Philosophical Analysis in Handbook of categorization and cognitive science

[19] (1978) Rosch E., Lloyd, B. B. Cognition and categorization. Hillsdale NJ: Erlbaum Associate

[20] Rosen, R. (1985). Anticipatory Systems, Robert Rosen, Pergamon Press

[21] Roy D. (2005) Semiotic schemas: a framework for grounding language in action and perception, Artificial Intelligence, Volume 167 , Issue 1-2, pp. 170-205

[22] Sowa, J. (2005) Categorization in Cognitive Computer Science in Handbook of categorization and cognitive science, 2005,

[23] Sowa, J. (2006) Worlds, Models, and Descriptions, Studia Logica, Special Issue Ways of Worlds II, 84:2,, pp. 323-360.

[24] Tsang, E. (1993). Foundations of Constraint Satisfaction. Academic Press.

[25] Wittgenstein, L. (1961) Tractatus Logico-Philosophicus, Routledge and Kegan Paul