# A real-time agent system perspective of meaning and sapience

Ricardo Sanz, Julita Bermejo, Ignacio López, Jaime Gómez

Autonomous Systems Laboratory, Universidad Politécnica de Madrid, Spain,
ricardo.sanz@upm.es

## 1 Introduction

Wisdom and sapience have been traditionally considered desirable traits in humans, but the use of the terms is decaying perhaps due to a raising post-modern relativism that lessens the value of others' knowledge. This chapter proposes an interpretation of *sapience* in terms of meaning generation and knowledge exploitation in social groups of knowledge-based agents.

We will describe *sapient agents* as those that are able to generate useful meanings for other agents beyond their own capability of generation of self-meanings. This makes sapient agents specially valuable entities in agent societies because they provide interagent reliable third-person meaning generation that provides some functional redundancy that contributes to enhance individual and social robustness and global performance. This approach to meaning generation is pursued by our research group in the context os the ASys Theory of autonomous cognitive systems.

Knowledge-based systems have been a matter of research and development from years; from the logic-based problem solvers of the sixties to the expert systems of the eighties or contemporary model-based systems, the nature of exploitable knowledge has been a core issue in artificial intelligence. Construction of well performing systems seems to require the codification of suitable knowledge in suitable forms for the agent activity.

In a sense, there has been a raising awareness that having knowledge —whatever its form— is not enough. To perform adequately, agents need to acquire an *understanding* of their action context so they can rationally decide about the proper action to be taken and the proper knowledge to be used in deciding about it. This means that agents should interpret information coming from their sensors and generate meanings from this information to be used in the action decision-making process. This issue of *situation awareness* has been raised many times and even addressed specifically in the design of intelligent system architectures (see for example Figure 1).

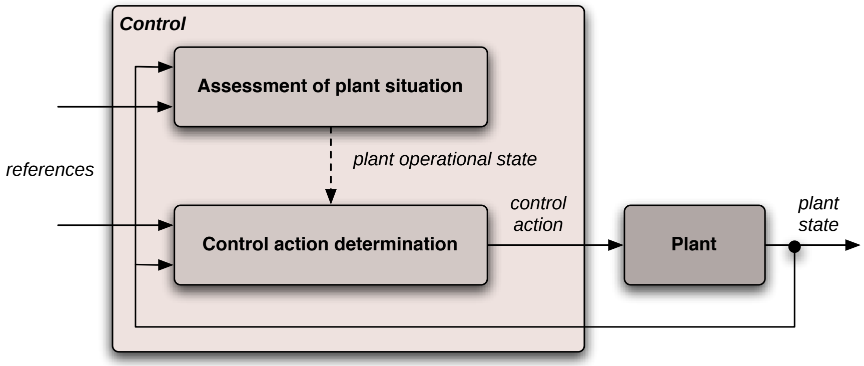


**Fig. 1.** Two level phasing of situated intelligent systems: 1) plant situation awareness and 2) control action generation; from Sanz (1990).

While this brief analysis directly enters the old debate about data, information, knowledge and meaning, we will not contribute extensively to it; but it will be somehow necessary to clarify some of the terms used in the analysis of wisdom and sapience that follows (e.g. intelligence, meaning or knowledge).

Mayorga (2005) proposed a differentiation between *intelligence* and *wisdom* based on the inner architecture of action. He sees ”Intelligence” as related to an ”Analysis” → ”Action” process; whereas, ”Wisdom” is seen as related to ”Analysis”, ”Synthesis”, → ”Action” process.

Although we are not going to enter the debate about the definition of intelligence (see (Sanz et al., 2000) for a partial account of our views that we can summarize as *utility maximisation in knowledge-based action*) it may be necessary to analyze the nature of the knowledge involved in action generation and propose a model for third-person meaning generation that will provide a simple interpretation of the concepts of ”wisdom” and ”sapience”.

To achieve this objective, firstly we will present a model of first-person meaning generation. Next, we apply this model to a cross-agent meaning generation process.

Other authors (Tien, 2003) consider that *wisdom* is just a further step in the *data* → *information* → *knowledge* ladder (see Figure 2). Or as Landauer puts it in his *meaning hierarchy*, the ladder is *data* → *information* → *knowledge* → *understanding* (Landauer, 1998).

While meaning (semantics) is critical for purposeful action, few psychological theories of mind have taken the study of meaning as the foundation of a working theory of the mind (Combs, 2000).

Hardy (1998) says that the generation of meaning is produced by the continuous closed causal link between an internal context (what she calls the semantic constellations), and an external context (a meaningful environment).

Other's argue for a theory of meaning based on embodiment. This alternative is based on the idea of embodiment (e.g., Barsalou, 1993; Glenberg, 1997; Lakoff, 1987),

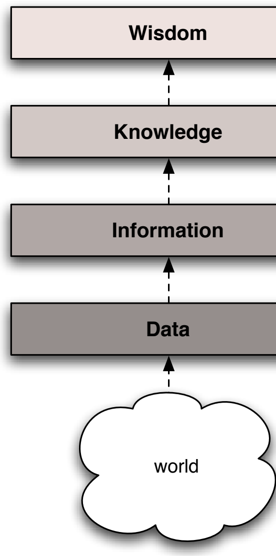


**Fig. 2.** Moving from information to wisdom according to Tien (2003).

that cognition is intimately connected with the functioning of the body (Glenberg et al., 1999).

## 2 The nature of meaning

Beyond classical accounts of life-related information and meaning generation (Oyama, 1985), we will focus on abstract cognitive agents with the –perhaps hopeless– purpose of having a theory applicable both to the analysis of extant cognitive agents and also to engineering processes of high-performance artificial agents, as those found controlling the technical systems of today's world.

Some authors have proposed that *meaning* is just a list of features —like a frame in classical AI— but there are compelling arguments from different sources against this interpretation (see for example (Shanon, 1988)). Another classic alternative was to consider that the meaning of symbols is a semantic network; but this leads to a recursive search of meaning that finally ends in the symbol grounding problem (Harnad, 1990). A third solution is based on the symbols taking on meaning by referring to entities outside the agent. That is, perception is seen as the core engine of meaning assignment to internal symbols. This corresponds to the views of interactivist schools; but the recurrent discussion about the necessity of embodiment will disappear when constructors become aware that minds necessarily run on virtual machines and hence the existence and awareness of an extant body is both unavoidable and useful for enhancing behavior.

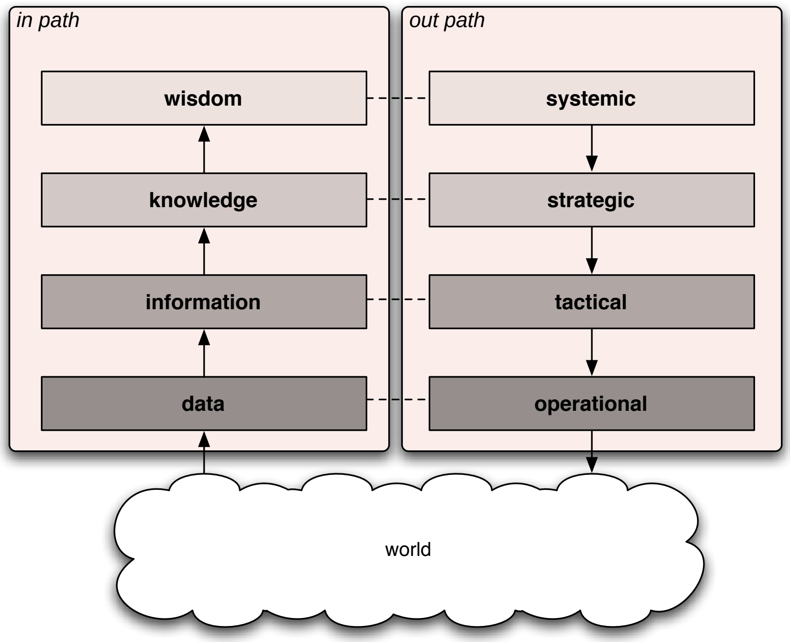


**Fig. 3.** The *in* and *out* paths of a situated system show the range of decision-making activities coupled with the different *information* levels.

In most of these interpretations, however, there is a big pending issue; they usually lack support of a core feature of meanings: *meanings can capture the dynamics of entities in their contexts*. Meanings are not constrained to statics but do also express change (actual or potential).

If we can say that X captures the meaning of a concrete piece of information it is because X provides a sensible account of the relation of the agent with the originator —the causal agent— of the information in present and potentially future conditions.

As Meystel (2001) says, "the first fundamental property of intelligent systems architectures (the property of the existence of intelligence), can be visualized in the law of *forming the loop of closure*" (See Figure 4). This loop of closure is seen in intelligent systems as composed by the world, sensors, world models and behavior generators, this last three constituting parts of the agent. A fourth component is necessary to provide the goal-centered behavior of agents: the value judgment engine.

If we consider how behavior is generated, the value judgment component in RCS architecture is critical (see Figure 7). But this value judgment shouldn't be done over raw or filtered sensor data (i.e. judging the present state of affairs) nor over the agent's present mental state. Value judgment is necessarily done over potential futures based on agent't present mental state. Is this value judgment of potential future states what assigns meanings to facts of reality.

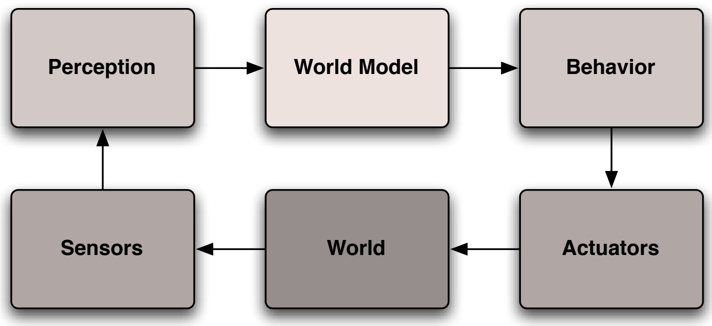


**Fig. 4.** The elementary loop of functioning —*loop of closure*— as described by Meystel (2003).

## 3 Meaning generation in the ASys model

The previous analysis shows that the core elements of a meaning generation engine are a predictor and a state value calculator. This is what our brain does all the time to generate meanings: evaluation of causal impact of what we see. Meanings are generated by means of temporal utility functions.

A real-time time/utility function expresses the utility to the system of an action completion as a function of its completion time. This is seated at the core root of real-time systems engineering, i.e. engineering of systems that have requirements related to the passage of time.

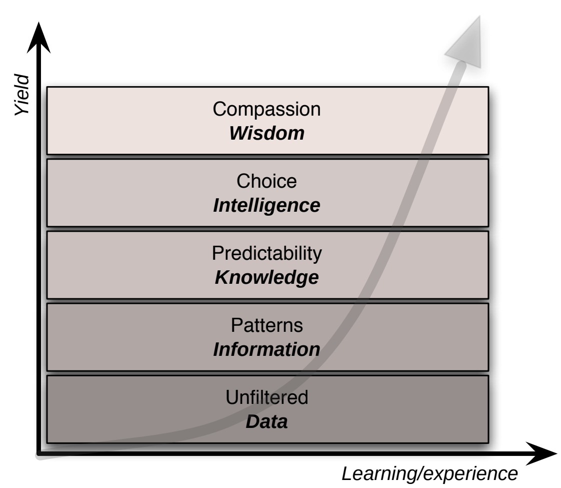


**Fig. 5.** The hierarchy of information/meaning as of Tuomi (1999) when an agent is evolving from the experiential viewpoint (yield = intellectual dividends per effort invested).

The meaning of a concrete perceived (externally or internally) fact is the partitioning of potential future trajectories of the agent in its state space. For example, if I see it's raining outside, this fact divides all my potential futures into two sets: in one I continue dry; in the other one I get wet. This partition *is the meaning of the fact* "it's raining outside".

This interpretation of meaning as related to the dynamics of futures can be found in many different areas, for example, neurobiology, psychology or even software engineering.

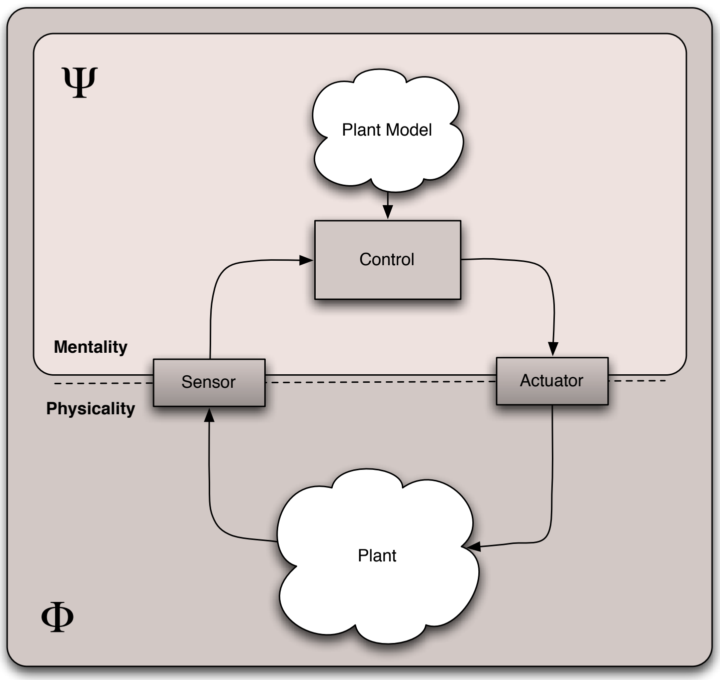


**Fig. 6.** Situated cognitive agents exploit the interaction with the world to maximise utility and this is achieved by means of driving such interaction by means of models of the reality —the plant under control in artificial systems— that constitute the very knowledge of the agent.

In order to help in this calculation of futures and future values, situated cognitive agents exploit the interaction with the world —and with other cognitive agents— to maximise behaviour utility by means of driving such interaction by adaptive models of the reality they are dealing with (see Figure 6).

These models of a part of the reality —the plant under control in artificial systems— constitute the core of the real world knowledge of the agent, and are the very foundation of meaning calculation.

# 4 Other analyses of meaning generation

## 4.1 Freeman's mental dynamics

Walter Freeman identifies *meanings* with "the focus of an activity pattern that occupies the entire available brain"(Freeman, 1997). From his point of view there are no representations in the brain, only meanings. The brain is an engine for meaning generation –based on brain perceptual dynamics– and, simultaneously, an engine for action generation based on the same type of dynamics.

## 4.2 Gibson's affordance theory

According to the ecological psychologist James Gibson (Gibson, 1979), an *affordance* is an activity that is made possible –an *action possibility* so to say– by some property of an object. A valve affords flow control, by being of the right shape and size and being in the proper pipe place where one needs to reduce flow.

In some contexts, affordances are classified into three categories: based on sensory (unlearned sensory experience), perceptual (learned categorizations of sensory experience) or cognitive (thought-based) processes. There are even considerations about the possibilities of non-aware affordances.

The most classic example of affordances involves doors and their handles (buildings, cars etc.) but the world of control systems is full of these entities: actuators are embodiments of affordances.

## 4.3 Griswold's programs meaning

In the area of tools-based software engineering, programmers look for automated methods of automated transformation of program specifications into final deployable packages. This is expect to solve the handcrafting bottleneck of manual programming. See for example, the work of Griswold and Notkin (1995) in the field of computer program transformation.

This implies having *meaningful transformations* of programs between different representations. The MDA proposal, for example, considers transformations from UML-based Platform Independent Models into platform-dependent models, and then into concrete implementation-oriented languages (IDL, C++, etc.).

All this transformations should be, however, meaning-preserving. But program meaning is not related with the actual wording of the code –that in model-centric software development may not even exist in some phases– but with the concrete program functionality (the program behavior) when executed over the appropriate platform, i.e. the platform that provides the required abstractions that the application was based upon.

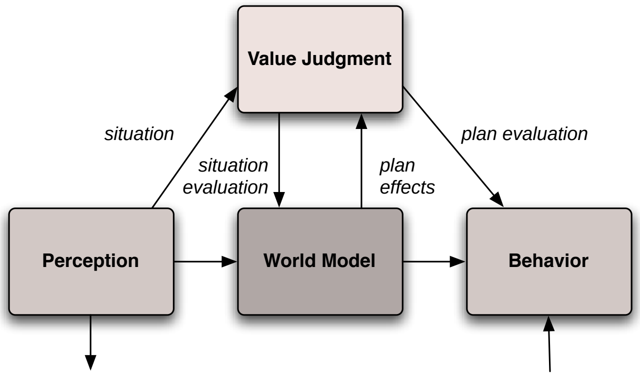


**Fig. 7.** The elementary loop of functioning of Meystel incremented with a *value judgment* unit to generate *meanings*; this design matches what is proposed in the ASys Theory about menaing generation. This structure corresponds to the elementary control node of the RCS intelligent control architecture (Albus, 1992).

## 5 Meaning in control systems

From the former analysis, we can see that meaning cannot be associated to an isolated piece of information but to a set composed by the *information*, the *agent* for which the information is meaningful and the *context* where the agent operates. To summarize, the meaning of a piece of information is agent- and context-dependent, something that it is well known in psychology (Clark, 1998).

Most researchers' creatures manipulate meanings without having an explicit theory of them; by means of ad-hoc meaning generation processes embedded in the control architectures. These are based on a particular, hidden ontology and a value system that is implicit in the architecture (see for example the work of Steels (1998)).

Valuable engineering efforts are those oriented toward a clarification of the role that architecture plays in control systems and how is it possible to attain constructabiliy of complex systems by means of scalable design patterns. This approach is specially well captured in the multiresolutional approach fostered by the control design pattern that Meystel calls the *elementary loop of functioning* (Meystel, 2003). Of importance in relation with the ASys theory of meaning is the incorporation of value judgment mechanisms over this elementary loop (see Figure 7).

The elementary loop of functioning, when applied hierarchically, generates a multiresolutional ladder of meanings specifically focused on the controllable subspace of each control level. This approach partitions both the problem of meaning generation and the problem of action determination, leading to hierarchical control structures that have interesting properties of self-similarity.

This core design pattern approach is extended in the concept of a control node of the RCS control architecture (Albus, 1992) (see Figure 7). Beyond the model of the world and the sensing and acting units, this architecture considers the existence of a value judgment unit that evaluate both static states and dynamic states derived from hypothetical plan execution.

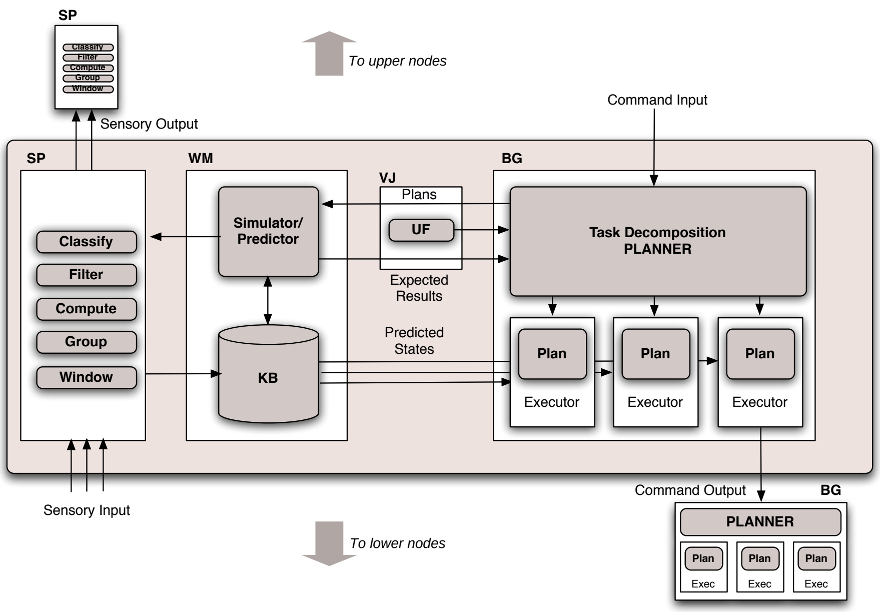


**Fig. 8.** The basic RCS node interchanges sensory and command flows with upper and lower nodes. While this may be considered *meaningful* flows, their meaning —*sensu stricto*— is limited to the originating node (Albus and Barbera, 2005).

## 6 Sapience: Generating other's meanings

To go to the core issue of the problem, i.e. the nature of sapience, we interpret it as the *capability of generating meanings for others*. Sapient agents can interpret the state of affairs and generate meanings that are valuable for other agents, i.e. like those generated by value judgment engines that are transpersonal. The attribution of sapience is *social* in the sense that it happens when the sapient agent is able to generate meanings that are socially valid, i.e. valid not only for one agent but for a group of agents. Generating meanings that are valid for more that one agent is beyond normal agent capabilities. That makes sapient agents really special.

To some extent, sapient systems can voluntarily select and use shared ontologies (that are used by others) and prediction engines to generate meanings that are valid for them. This capability of shared ontology selection and use is largely sought (Mizoguchi and Ikeda, 1996) in present-day research on distributed information systems (see for example the efforts related with the semantic web).

Beyond the meaning calculation fact, sapient systems do usually manifest themselves by means of their explanatory capabilities; i.e. they can communicate the results of the calculation to the target agent. This may be seen as clearly rejecting those fashionable accounts of sapience as obscure manifestations of mental capability. Explanation is hence strongly related with the perception of sapience (see (Craik, 1943), (Brewer et al., 1998) or (Wilson and Keil, 1998)).

Obviously this vision is strongly related with the psychology concept of "theories of mind" but goes well beyond it in the sense that the "theory of mind" is typically restricted to agent-to-agent interaction.

This view of sapience can be implicit or explicit (when the sapient system uses consciously the model of the other to calculate meanings). It is like having 'deliberative' sapience.

## 7 Meanings in hive minds

Of major interest for us, that focus our research in the domain of complex distributed controllers, is the capability of exploiting this sapience mechanics to improve the integration level of a distributed controller.

We may wonder to what extent meaning integration can lead to mind federation and the emergence of a single, unified controller: a hive mind. If meaning is globally integrated this implies that the different subsystems may be aware of what is going on affecting other subsystems. A kind of distributed consciousness emerges.

Some people have considered the possibility of shared or collective consciousness even for humans (see for example Hardy (1998), Sheldrake (1988) or Laszlo (1996)). From this perspective, individuals can conjointly share a particular experience even being at distance.

People dealing with practically independent environments, can use other's previous experiences in similar situations to better understand the present state of affairs. These previous experiences are culturally shared and when executed over similar virtual machines (Sloman and Chrisley, 2003) can generate similar interpretations of reality that coalesce into coherent social behaviors that can be seen as a form of collective understanding.

Perhaps we can exploit this kind of social phenomena in the implementation of advanced cognitive conscious modular controllers.

## 8 Conclusions

Agent's meanings are not static interpretations of agent-perceived data but do capture future trajectories of the agent in his state space in a particular context. This is strongly related to Putnam's causal theory of meaning (Putnam, 1975).

Sapient systems are agents that have the capability to generate meanings for others, i.e. they can assess situations as other agents would do and suggest courses of action based on other agents' set of values.

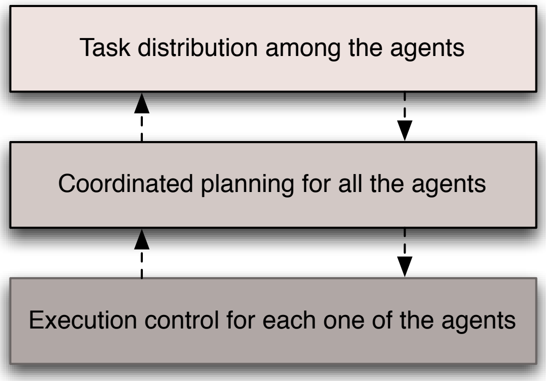


**Fig. 9.** Multi-agent systems can only operate if the ontologies are shared to be able to reconstruct meaning from messages coming from other agents.

Wisdom is hence nothing categorically different from what is available in conventional agent architectures but a particular capability of an agent to use its own resources to think-for-others. Wisdom in hence attributed by other's due to this capability that goes beyond usual agent capabilities.

This understanding of meaning is strongly related with recent theories of consciousness and lead us to the possibility of achieving consciousness states in control systems (Sanz and Meystel, 2002).

This approach to explicit management of meanings is currently under implementation in the SOUL Project (http://www.aslab.org/public/projects/SOUL/) in the laboratory of the authors.

## 9 Acknowledgements

The authors would like to acknowledge the support of the Spanish Ministry of Education and Science through the DPI C3 grant and the support of the European Commission through the IST ICEA grant.

## References

Albus, J. S. (1992). A reference model architecture for intelligent systems design. In Antsaklis, P. and Passino, K., editors, *An Introduction to Intelligent and Autonomous Control*, pages 57–64. Kluwer Academic Publishers, Boston, MA.

Albus, J. S. and Barbera, A. J. (2005). Rcs: A cognitive architecture for intelligent multi-agent systems. *Annual Reviews in Control*, 29(1):87–99.

Antsaklis, P. and Passino, K. E. (1993). *An Introduction to Intelligent and Autonomous Control*. Kluwer Academic Publishers.

Brewer, W. F., Chinn, C. A., and Samarapungavan, A. (1998). Explanation in scientists and children. *Minds and Machines*, 8:119–136.

Clark, A. (1998). Twisted tales: Causal complexity and cognitive scientific explanation. *Minds and Machines*, 8:79–99.

Combs, A. (2000). Book review: Networks of meaning: The bridge between mind and matter, by Christine Hardy. *Nonlinear Dynamics, Psychology, and Life Sciences, Vol. 4, No. 1, 2000*, 4(1):129–134.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press, London.

Freeman, W. J. (1997). A neurobiological interpretation of semiotics: meaning vs. representation. In *1997 IEEE International Conference on Systems, Man, and Cybernetics*, volume 2, pages 12–15.

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Houghton Mifflin, Boston.

Glenberg, A. M., Robertson, D. A., Jansen, J. L., and Johnson-Glenberg, M. C. (1999). Not propositions. *Journal of Cognitive Systems Research 1*, 1:19–33.

Griswold, W. and Notkin, D. (1995). Architectural tradeoffs for a meaning-preserving program restructuring tool. *IEEE Transactions on Software Engineering*, 21(4):275–287.

Hardy, C. (1998). *Networks of Meaning: The Bridge Between Mind and Matter*. Praeger/Greenwood Publishing Group, Westport, CT.

Harnad, S. (1990). The symbol grounding problem. *Physica D*, 42:335–346.

Landauer, C. (1998). Data, information, knowledge, understanding: computing up the meaning hierarchy. In *1998 IEEE International Conference on Systems, Man, and Cybernetics*, volume 3, pages 2255–2260.

Laszlo, E. (1996). *The whispering pond*. Element Books, Rockport, MA.

Mayorga, R. V. (2005). Towards computational sapience (wisdom): A paridigm for sapient (wise) systems. In *IEEE International Conference on Integration of Knowledge Intensive Multi-agent Systems*, Boston.

Meystel, A. (2001). Multiresolutional representation and behavior generation: How do they affect the performance of intelligent systems. In *Tutorial at ISIC'2001*, Mexico D.F.

Meystel, A. M. (2003). Multiresolutional hierarchical decission support systems. *IEEE Transactions on Systems, Man and Cybernetics*, 33(1):86–101.

Mizoguchi, R. and Ikeda, M. (1996). Towards ontology engineering. Technical Report AI-TR-96-1, The Institute of Scientific and Industrial Research, Osaka University.

Orden, G. C. V., Moreno, M. A., and Holden, J. G. (2003). A proper metaphysics for cognitive performance. *Nonlinear Dynamics, Psychology, and Life Sciences*, 7(1):49–60.

Oyama, S. (1985). *The Ontogeny of Information: Developmental Systems and Evolution*. Cambridge University Press, Cambridge.

Pustejovsky, J. (1990). Perceptual semantics: the construction of meaning in artificial devices. In *5th IEEE International Symposium on Intelligent Control*, volume 1, pages 86–91.

Putnam, H. (1975). *Mind, Language and Reality*. Number 2 in Philosophical Papers. Cambridge University Press, Cambridge.

Sanz, R. (1990). *Arquitectura de Control Inteligente de Procesos*. PhD thesis, Universidad Politécnica de Madrid.

Sanz, R., MatÌa, F., and Gal·n, S. (2000). Fridges, elephants and the meaning of autonomy and intelligence. In *IEEE International Symposium on Intelligent Control, ISIC'2000*, Patras, Greece.

Sanz, R. and Meystel, A. (2002). Modeling, self and consciousness: Further perspectives of ai research. In *Proceedings of PerMIS '02, Performance Metrics for Intelligent Systems Workshop*, Gaithersburg (MD), USA.

Shanon, B. (1988). Semantic representation of meaning: A critique. *Psychological Bulletin*, 104:7–83.

Sheldrake, R. (1988). *The presence of the past*. Random House, New York.

Sloman, A. and Chrisley, R. (2003). Virtual machines and consciousness. *Journal of Consciosness Studies*, 10(4-5):133–172.

Steels, L. (1998). The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103:133–156.

Tien, J. M. (2003). Toward a decision informatics paradigm: A real-time, information-based approach to decision making. *IEEE Transactions On Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 33(1):102–112.

Tuomi, I. (1999). Data is more than knowledge. In *Proceedings of the 32nd Hawaii International Conference on System Sciences*.

von Uexküll, J. (1982). The theory of meaning. *Semiotica*, 42(1):25–82.

Wilson, R. A. and Keil, F. (1998). The shadows and shallows of explanation. *Minds and Machines*, 8:137–159.

Ziemke, T. (2002). On the epigenesis of meaning in robots and organisms. *Sign Systems Studies*, 30:101–111.

Ziemke, T. and Sharkey, N. E. (2001). A stroll through the worlds of robots and animals: Applying Jakob von Uexk¸ll's theory of meaning to adaptive robots and artificial life. *Semiotica*, 134(1-4):701–746.

Zlatev, J. (2001). The epigenesis of meaning in human beings, and possibly in robots. *Minds and Machines*, 11.