Joslyn, Cliff A: (2000) "Virtual Environments as Constraints on Decision-Making in Agent Models of Socio-Technical Organizations", in: 2000 Workshop on Virtual Worlds in Simulation, Simulation Series, v. 32:2, ed. K. Bellman and C. Landauer, pp. 120-126, Society for Computer Simulation

# VIRTUAL ENVIRONMENTS AS CONSTRAINTS ON DECISION MAKING IN AGENT MODELS OF HUMAN ORGANIZATIONS

**Dr. Cliff Joslyn**

Computer Research Applications Group (CIC-3)
MS B265, Los Alamos National Laboratory
Los Alamos, NM 87545
joslyn@lanl.gov, www.c3.lanl.gov/~joslyn

**Keywords:** Agent-based modeling, semiotics.

## Abstract

We consider agent-based modeling from a semiotics perspective, with application to the simulation of Socio-Technical Organizations. After outlining our semiotic approach, we consider the role of constraints present in the virtual environment, including the virtual physics, common communication structures, and shared knowledge, for constraining the decision-making capabilities of the constituent agents.

### Introduction

The modern environment is an interlocking collective of large numbers of groups of people interacting with computer systems, and which themselves interact with a variety of physical systems to maintain them under conditions of good control. The vast complexity and quantity of information involved makes simulation approaches necessary, and yet the existing formalisms available for simulation are not sufficient to reflect their full characteristics. In particular, simulations built on strict formalisms such as discrete-event systems or hybrid control cannot capture the inherent freedom available to humans interacting with such systems; and simulations built on classical rule-based Artificial Intelligence (AI) approaches are too brittle and specific to allow for the emergent phenomena which characterize such systems.

Recently, emphasis has been placed on concepts of autonomous and/or intelligent *agents* as the hallmark of a new paradigm for information systems. An agent-based modeling approach between collective automata systems such as used in Artificial Life (ALife) and full AI may provide a robust capability to simulate human-machine interaction at the collective level.

Our concept of a *semiotic agent* (Joslyn and Rocha 2000) distinguishes agents specifically as *decision-making* systems. These have a sufficient freedom over a variety of possible actions to make specific predictions of their actions impossible at the targeted scale of observation. Clearly this class includes AI systems, but leaves out simpler collective automata or state-transition systems typical of ALife. We call this approach semiotic, as it focuses on the use and communication of symbols by and between agents and their environments.

We have also argued (Rocha and Joslyn 1998) that interesting emergent behavior in agent systems must arise from considering them as systems in interaction with some form of environment with a sufficiently rich set of properties in and of itself. We thereby further distinguish semiotic agents from pure decision-making algorithms (Wolpert *et al.* 1999), in that they are embedded in (hopefully rich) virtual environments in which they take actions which have consequences for the future of the agents themselves. Thereby, these environmental interactions induce constraints on the freedom of decision-making on the part of the semiotic agents.

In this paper, we first discuss the current state of the use of agents with respect to both decision theory and modeling and simulation. We then outline the semiotic approach to agent simulation and our project in modeling Socio-Technical Organizations (STOs). We conclude with a discussion of recent work which demonstrates the significance of constraints present in the virtual enironments in agent-based simulations for greatly increasing system performance and/or robustness, and in particular those revealed by the decomposition of the virtual environment into a virtual physics, shared communication structures, and shared knowledge.

## Agent Approaches to Modeling and Simulation

The history of computer science has seen a "march of paradigms", as programming theory has moved from procedural through functional to object-oriented models, now culminating in this agent-based approach (Dyer 1999). Hype has now led to the situation where nearly anything can be identified as an agent: in robotics, robots are mechanical agents in real environments; in information systems agents are software servants of a user; in software engineering, agents are "super-objects", combining encapsulation with autonomous process control and indpendent threading; in ALife, agents are simple state-determined automata connected according to various complex temporal or topological schemes in order to demonstrate "emergent behavior" as complex dynamical processes; in AI, complex software system with a great deal of on-board computational intelligence and planning ability interact in small collections and simple environments; and finally in decision theory, political scientists regard individuals engaging in collective choice processes as agents.

In abstracting away from these disparate senses of agency we discern their common properties, including asynchrony, interactivity, mobility, distribution, and randomness of trials over various initial conditions. While this list is common to most agent models, they still do not capture the essential qualities which most people bring to the concept of "agency". These qualities are a kind of *independence*, the fact that the agent is doing something of and by itself. This refers to a kind of *self-control*, or, in a word, *autonomy*. From this property alone, all of the above follow.

We also recognize a number of aspects of the concept of autonomy, including boundaries distinguishing the domain of the agent from the rest of the environment; the quantitative degree of autonomy, since this is a relative concept; the identity of the autonomous system which follows from the identification of the boundary, and thus the ability to distinguish between that which is inside and outside the boundary; and finally a closure of some aspects of the world which are entrained within the boundary (identity) of the agent and thereby closed off from other interactions (Joslyn 1998).

The list of properties above is actually quite familiar to us from the foundations of systems theory (Bunge 1992, Joslyn 1999). Indeed, based on the above criteria there is very little to distinguish an "agent" from some general sense of "system". In seeking a coherent sense of agent that will be distinguished not only from other software engineering senses (agents are not just subroutines or objects), but also from "objects" or "systems" in general (agents are not just systems), we focus on the concept of **autonomy with respect to action**. In other words, our concept of agent is a system (object) which has an inherent *freedom* to make *choices* or *decisions* over possible *actions*.

## Semiotic Agents

We will call such agents *semiotic* to distinguish them from all others. We will discuss the particular sense of semiotic system below, but say here that it is related to some common ideas in the literature, particularly the AI concepts of "reactive" vs. "deliberative" processes (Sloman and Brian 1999), and indeed, we would argue that all (deliberative) AI systems are semiotic in that sense. However, we are also motivated by the ALife and complex systems critique of AI, which allows for emergent phenomena through autonomy as opposed to external programming of elaborate internal models and planning mechanisms. Thus our goal is to construct semiotic agents which are sufficiently, but *minimally* sufficiently, complex to have autonomy of action.

Comparing traditional ALife approaches of large collections of simple agents with the AI approach of small collections of complex agents, our goal is to aim solidly between them, implementing agents which are relatively simple, and thus whose collections can have emergent properties, but with sufficient memory bases and uncertainty structures to allow for deliberative capabilities.

Our fundamental architecture for semiotic agents is shown in Fig. 1. The system takes measurements from its environment, and constructs generalized "beliefs": stored representations of the current state and memories of past states. There is also an internal representation of "desires", namely potential goals states. A decision node decides among potential actions, which are then taken back into the environment. Finally, those actions interact with the dynamical processes in the environment, which then feed back to the agent in the form of future perceptions. In this way, the consequences of the agent decisions have an explicit impact on its future development.

This architecture is based on the principles of a generalized control architecture, where the autonomy of the system is allowed by its manifestation of a closed causal generalized negative feedback control relation with its environment (Powers 1989); the autonomy of action necessary for semioic agents is al-

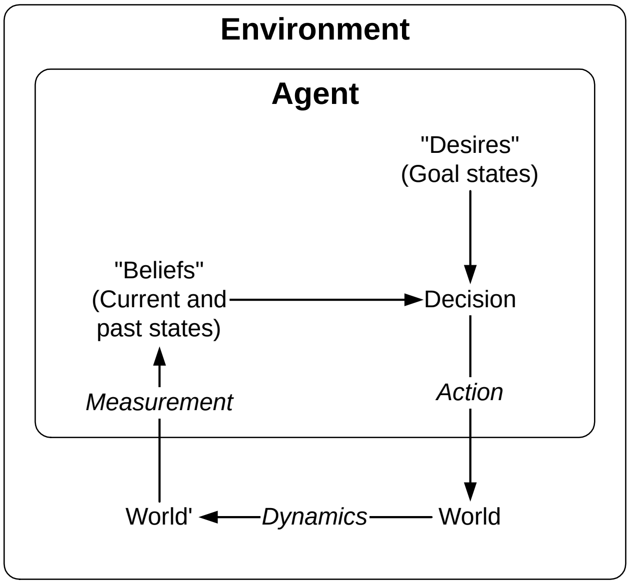


Figure 1: Architecture of a semiotic agent.

lowed in virtue of the dynamical incoherence of the memory structure and the indepenent representation of the decision function; and in avoiding AI implementations, casting beliefs and desires as relatively simple, non-propositional uncertainty structures.

What characterizes these systems is that they involve processes of perception, interpretation, decision, and action with their environments. The memory structures required by such dynamically incoherent systems further entails the presence of *representations* stored internally to the agent, in particular of measured states of affairs, goals, and possible actions.

Thus we turn to semiotics, or the general science of signs and symbols. Originally a humanities subfield of linguistics (Deely 1990), semiotics has come to become more prominent first in text and media analysis, and then in biology (Deely 1992), computer engineering, and control engineering (Meystal 1996).

Semiotic processes involve the interpretation of sign tokens maintained in coding relations with their interpretants. Thus semiotics in general is concerned with issues of sign typologies, digital/analog and symbolic/iconic representations, the "motivation" (intrinsic relations of sign to meaning) of signs, and mappings among representational systems.

Semiotics further decomposes semiotic relations along three axes: syntactic, concerning the formal relations among sign tokens; semantic, concerning the interpretation of tokens by agents as standing for environmental observables; and pragmatic, concerning the repurcusions of those interpretations for the agent in its environmental context, in other words, the purposes or goals of sign interpretation.

Thus the semiotic agent-based modeling approach can be summarized as follows:

- Simulated agents operate within **environments** which *have their own rules*, or their own "virtual physics".
- Agents have **action capabilities** which must be considered *relative to those environments*.
- The possible **decisions** that agents can make must be considered *relative to those possible actions*. Thus we assert that pure decision models such as (Richards, McKay and Whitman, 1998; Wolpert *et al.*, 1999) cannot fully realize the full emergent capabilities of agent communities in complex environments. Various decision capabilities include deterministic input/output state systems, mutable transfer functions in terms of evolutionary (external selection) or adaptive (internal selection) processes, and finally, the use of culture as shared knowledge among agents to aid in agents decision-making.
- **Data** is seen as *information transmission among agents*.
- **Knowledge** is seen as the *interpretation of data* by agents
- **Communication** among agents must be seen as *relative to the knowledge and internals* of the agent.
- **Control** is seen as a form of *decentralized constraint over decision-making* in agents, potentially from many sources, including everything above.

## Modeling Socio-Technical Organizations

Our application area is in the simulation of **socio-technical organizations** (STOs) such as generalized command and control organizations or utility infrastructure systems. The pressing needs are to assess the stability and vulnerabilities of STOs, and to protect their robustness against disruption in the event of destablizing forces, such as inherent dynamical instability, structural modification, or information disruption or disinformation, perhaps through deliberate attack or sabatouge.

STOs are characterized by a complex structure involving the hybrid interaction of physical systems with agent (human) organization Fig. 2:

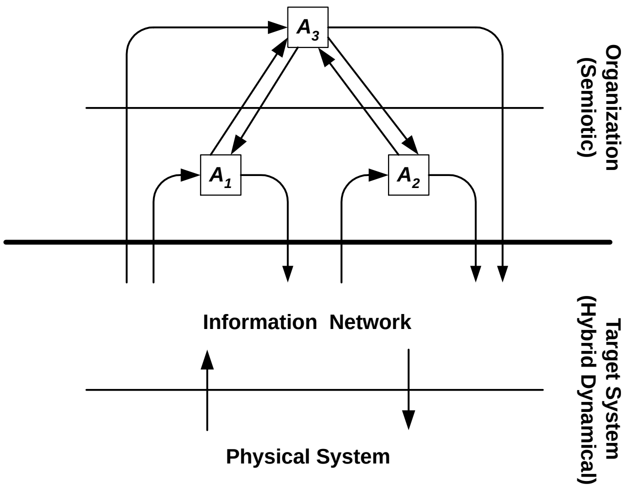


Figure 2: Socio-Technical Organizations.

- A **target system**, which itself consists of: a **physical system** which is deterministic (typically, and as we will assume here, a continuous dynamical system), involving the flow of physical objects or substances through a complex enviornment ("terrain"); and an **information network** which is semi-automated, largely computer-based, and dependent on data acquisition, telemetry information, and control actions with the dynamical system.

- The target sytem acts as the environment to an **organization** of (human or computationl) agents or actors, which also has a complex structure consisting of: **operators**, atomic units which interact in prescribed ways with the information network; **supervisory** levels, which establish operational boundaries over lower or parallel systems, and alter system parameters; and ultimately the goals of the various corporate, military, and/or governmental organizations involved, including economic and political forces.

The boundaries among these levels, in particular between the target system and the organization, should be drawn *functionally*, distinguishing those components which *must* be considered as semiotic agents in virtue of their autonomy of decision making, and those which might not be. In particular, a human, if sufficiently constrained by conditions in the environment or communication system, might be representable as a deterministic component of the target system; and conversely, a computer system of sufficient freedom and complexity might be considered part of the organization.

## Bounded Freedom on Decision Making

In STOs, the agents are culturally coherent, with a broad base of shared knowledge; agent communication and interaction is mediated by a computer-based information network with a specific set of protocols; and agents must take actions in a real physical environment (terrain or an infrastructure network).

These three aspects of shared knowledge, shared communication structure, and a shared virtual physics thus compose the virtual environments in agent simulations of STOs. In turn, each of these aspects is also addressable from a semiotic perspective, by emphasizing that semiotic agents operate in a context of *bounded freedom* on their decision-making capacities.

Typical agent models possess one or more of these aspects, but they all become necessary when developing agent models of STOs. Recently, researchers have demonstrated that such constraints can be crucial in providing robustness and stability in multi-agent systems. In the remainder of this paper, we will point to some recent results from the literature which explicate this.

### Virtual Physics

Agents are embedded in a virtual physical environment, whether simulating aspects of a real environment or a purely synthetic world. Decisions about actions are thereby constrained relative to the properties of these environments. Such constraint are a source of robustness and improved performance. Two recent research results provide examples.

Gordon and Spears (Gordon *et al.*, 1999; Spears and Gordon, 1999) use agent models to simulate distributed sensor grids. Their virtual environment consists of a two dimensional continuous square grid, in which point agents with masses $m_i$ are embedded. The environment is equipped with a virtual force law $F = Gm_im_j/r^2$, where $r$ is the distance between any two agents, and an arbitrary distance constant $R$. When $r > R$ the force is attractive, and when $r < R$ repulsive. When $r > 1.5R$ there is no effect.

An agent's only possible action is to move in response to the sum of the forces acting on it. In virtue of the geometric property of hexgonal structures arising from intersecting circles (Fig. 3), the result is the agent collective constructing an hexagonal grid as the system reaches equilbrium, as shown in Fig. 4.

In this case, the agents have a single property (mass) and action (movement). When granted a second property, a binary "spin", and a modification of the force law so that the distance $r$ is renomralization to $r/\sqrt{2}$ only if the two particles in question have the

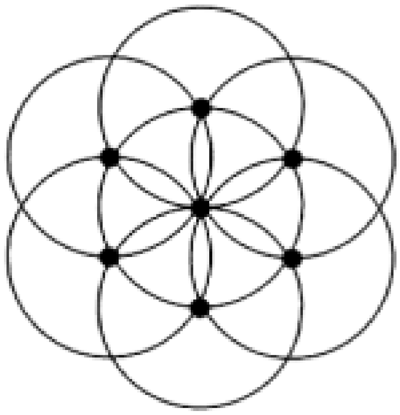

Figure 3: From (Spears and Gordon, 1999): "How circles create hexagons."

same spin, what results is a *square* lattice (see Fig. 5, Fig. 6).

Another example is provided in the work of (Pepper and Smuts 1999). Their experiment involves simulated evolution on a discrete square grid with various agents, including cells with vegetation, foragers, and predators. They were attempting to determine the conditions under which populations with stable mixtures of altruistic and selfish behaviors can arise.

The particular aspect of their simulation of interest to us here is the dependence on the structure of the virtual environment, in particular on the distribution of vegatation in the environment, which is a food resource for the foragers. Vegatation was distributed in square patches separated by gaps. Fig. 7 shows a patch size of three and a gap wideth of two.

Fig. 8 shows the distribution of the frequency of alarm-calling (an altrustic behavior) in the population as a function of patch width and gap. It is clear that an *intermediate* degree of structure in the environment is necessary to achieve robust performance: if the patches are too large or close together, sufficient resources are available to allow altruism to be swamped out as deselective; if too small or far apart, there are insufficient resources to support any population.

**Communication**

Agents can take actions into their environments, and can communicate with other agents. So an important question for any agent-based simulation is the following: is a communication act an actual action into the environment, or not?

Now it is certainly true that all communication acts through any real or simulation environment in

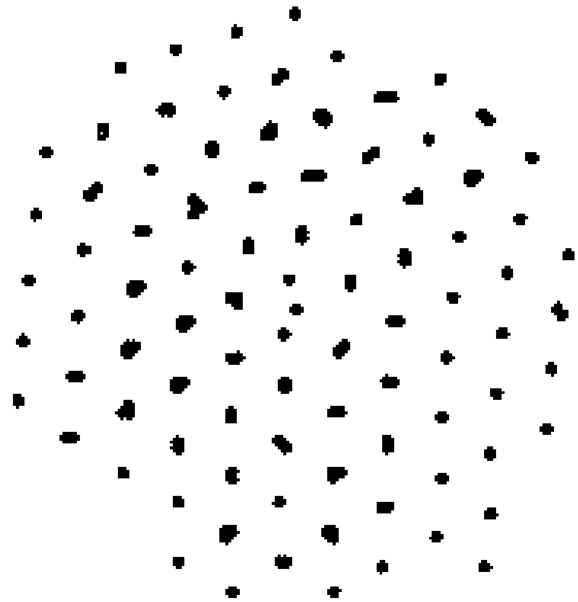

Figure 4: From (Spears and Gordon, 1999): "A good hexagonal lattice results by $t = 1000$."

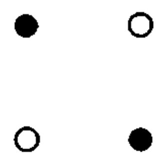

Figure 5: From (Spears and Gordon 1999): "Forming a square with particles of two spins."

fact take place through some kind of process in that environment. In real situations, our communication tokens have energy, require time to produce and transmit, usually have mass, and interact dynamically with physical processes. However, we usually ignore the energetic costs of token production, transmission, and reception, such as ignoring the thermal effects of speech or the mass of ink on paper.

Thus we will assume that in our agent simulations, communication is a distinct process, and is not part of the stream of actions taken specifically by agents. Thus the freedom of decision making of semiotic agents is also constrained by the semiotic structures used to record, transmit, and interpret information.

Perhaps the best example here is the long-standing semiotic work of Edwin Hutchins (Hutchins 1996, Hutchins and Hazlehurst, 1991). In both real naval systems and agent simulations of communication processes, Hutchins and his colleagues have explored or simulated agent learning mediated by situ-

| | Gap width | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Patch width | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 0 | * | * | * | * | * | * | * | * | * |
| 2 | 0 | * | * | * | * | * | * | * | * | * |
| 3 | 0 | 0 | 0 | 0.3 | **1** | **1** | **1** | **1** | **1** | **1** |
| 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0.6** | **0.6** | **0.8** |
| 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0.2 |
| 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.1 | 0 |
| 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Figure 8: From (Pepper and Smuts 1999), Table 2: "Final frequency of alarm-callers as a functin of path and gap width. * indicates extension."

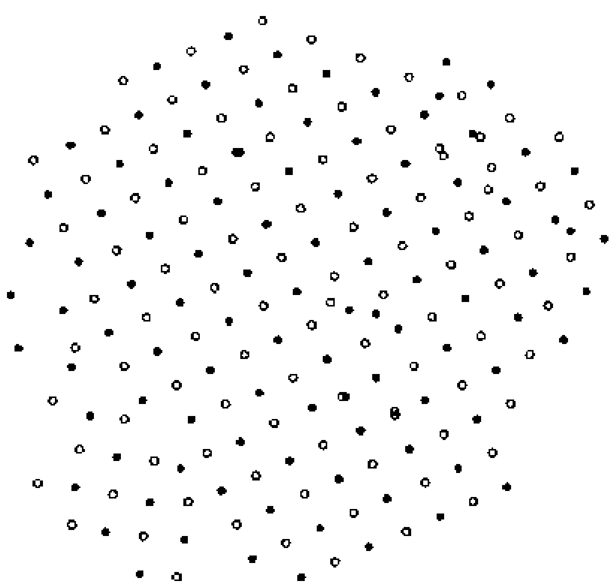

Figure 6: From (Spears and Gordon 1999): A square lattice (also uses a spin-flip repair mechanism).

ated exchange of semiotic tokens.

Hutchins and Hazlehurst (1991) simulated the ability of a community of evolutionary agents to learn environmental correlations, namely those between moon phase and tides (see Fig. 9). These neural-net based agents, in addition to standard learning, also generated communicational artifacts recording their knowledge, which were left in the environment to be discovered and interpreted by future generations of agents. Fig. 10 shows a "perfect" such artifact. In the experiment, learning mediated through such exchanges performed dramatically better than direct learning alone.

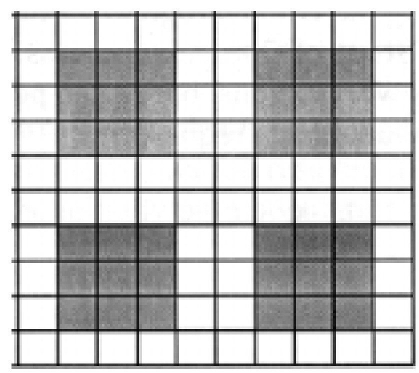

Figure 7: From (Pepper and Smuts 1999), Figure 1: "A representative resource distribution pattern."

### Shared Knowledge

Finally, decisions of agents may be constrained by a shared set of knowledge or beliefs, for example through a common biological evolution or cultural transmission (training or education). A cogent example here is the work of Richards, McKay and Richards (1998).

They modified the decision-theory problem of collective choice on the part of a group of agents with mutually conflicting preferences. Arrow's classical result shows that non-transitive collective behavior can result in cycling among the choices, and thus no overall result. Richards *et al.* have demonstrated that if the possible preferences are supplemented with shared knowledge among the agents in the form of structural relations among the choices (e.g. $A$ is directly related to $B$ and $C$, but $C$ is not related to $B$ directly), then unique collective choices are available for sizes of decision sets much larger than without.

TABLE 1 Citizen Language

| Environment | Symbolic Rep | Physical Rep |
|---|---|---|
| | Prototypic Moon Lexicon | Moon Phase |
| New moon | "1000" | 00 |
| First quarter | "0100" | 10 |
| Full moon | "0010" | 11 |
| Third quarter | "0001" | 01 |
| | Prototypic Tide Lexicon | Tide State |
| Large-variance tide | "01" | 0 |
| Small-variance tide | "10" | 1 |

Figure 9: From (Hutchins and Hazlehurst 1991), Table 1: Moon and tide "language".

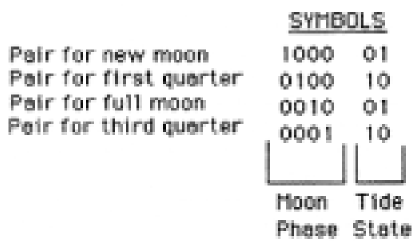


Figure 10: From (Hutchins and Hazlehurst 1991), Figure 4: "A perfect artifact".

## Acknowledgments

Much of this work was carried out within a research project on agent-based modeling of socio-technical organizations with Luis Rocha, who has contributed greatly to the exploration and explication of many of these concepts.

# References

[1] Bunge, Mario: (1992) "System Boundary", *Int. J. of General Systems*, v. **20**, pp. 215-219

[2] Deely, John: (1990) *Basics of Semiotics*, Indiana UP, Bloomington IN

[3] Deely, John: (1992) "Semiotics and Biosemiotics: Are Sign-Science and Life-Science Coextensive?", in: *Biosemiotics: The Semiotic Web 1991*, ed. TA Sebeok, J Umiker-Sebeok, pp. 46-75, Mouton de Gruyter, Berlin/NY

[4] Dyer, Douglas E: (1999) "Multiagent Systems and DARPA", *Communications of the ACM*, v. **42**:3, pp. 53

[5] Gordon, Diana; Spears, William; and Sokolsky, O et al.: (1999) "Distributed Spatial Control, Global Monitoring and Steering of Mobile Physical Agents", in: *Proc. 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA 99)*

[6] Hutchins, Edwin: (1996) *Cognition in the Wild*, MIT Press

[7] Hutchins, Edwin and Hazlehurst, B: (1991) "Learning in the Cultural Process", in: *Artificial Life II*, ed. C. Langton et al., MIT Press

[8] Joslyn, Cliff: (1998) "Models, Controls, and Levels of Semiotic Autonomy", in: *Proc. 1998 Conference on Intelligent Systems*, ed. J. Albus and A. Meystel, pp. 747-752, IEEE, Gaithersburg MD

[9] Joslyn, Cliff: (1999) "Levels of Control and Closure in Complex Semiotic Systems", in: *7th Annual Washington Evolutionary Systems Conference*, ed. G. van de Vijver et al.

[10] Joslyn, Cliff and Rocha, Luis M.: (2000) *A Semiotic Approach to Agent-Based Modeling*, in preparation for *Society for Computer Simulation Transactions on Simulation*

[11] Meystel, Alex: (1996) "Intelligent Systems: A Semiotic Perspective", *Int. J. Intelligent Control and Systems*, v. **1**, pp. 31-57

[12] Pepper, John W and Smuts, Barbara B: (1999) "The Evolution of Cooperation in an Ecological Context: An Agent-Based Approach", in: *Dynamics of Human and Primate Societies: Agent-Based Modeling of Social and Spatial Processes*, ed. TA Kohler and GJ Gumerman, Oxford UP, New York

[13] Powers, WT, ed.: (1989) *Living Control Systems*, CSG Press

[14] Richards, Diana; McKay, Brendan D; and Richards, Whitman A: (1998) "Collective Choice and Mutual Knowledge Structures", *Advances in Complex Systems*, v. **1**:2-3, pp. 221-236

[15] Rocha, Luis M and Joslyn, Cliff: (1998) "Simulations of Evolving Embodied Semiosis: Emergent Semantics in Artificial Environments", in: *Proc. 1998 Conf. on Virtual Worlds in Simulation*, pp. 233-238, Society Comp. Sim., San Diego

[16] Sloman, Aaron and Logan, Brian: (1999) "Building Cognitively Rich Agents Using the Sim_Agent Toolkit", *Communications of the ACM*, v. **42**:3, pp. 71-77

[17] Spears, William M and Gordon, Diana F: (1999) "Using Artificial Physics to Control Agents", in: *Proc. 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA 99)*, reviewed

[18] Wolpert, David; Wheeler, Kevin R; and Tumer, Kagan: (1999) "General Principles of Learning-Based Multi-Agent Systems", in: *Proc. 3rd Int. Conf. on Autonomous Agents*, `http://ic.arc.nasa.gov/ic/people/kagan/pubs/agents99_bar.ps.`