EXTENDING SITUATED REASONING: A MODEL AND A MECHANISM

David L. Sallach
University of Chicago
sallach@uchicago.edu

In synthesizing a hybrid multiagent model, Muller (1996, pp. 17-43) identifies three levels of agent response: reactive, deliberative, and interactive, and implements a layer for each type of response. Although this type of synthesis advances agent models by integrating diverse response patterns within a single agent, it also faces the difficulty of integrating multiple layers of relatively discrete mechanisms (i.e., situated-action rules, modal logic, coordination protocols). The present paper addresses this issue by using an affectivity heuristic, implemented as a proximity attractor, to provide an extended approach to situated reasoning.

Affectivity Heuristic. The agent model underlying the present proposal is drawn from empirical investigations of mass belief systems. Brady and Sniderman (1985), for example, observe that members of mass publics accurately estimate the political beliefs of strategic groups, and the policy positions group members are likely to take on salient issues. Such political insights exist in spite of the fact that mass publics possess a low level of abstraction and little specific information.

The mystery may be explained by their identification of an affectivity heuristic in which respondents combine their representations with an emotional reaction to the group in question. This response may be based upon personal experience, a shared generalization, or some other social basis. The affectivity heuristic appears to serve as an intellectual shortcut that allows the members of the public to draw inferences from complex domains of varying salience.

Nothing suggests, however, that the cognitive process on which these effects are based forms a special-purpose heuristic applicable only to public belief. On the contrary, the present discussion suggests that the dynamics of mass belief systems provide a model for the integration of multiple agent levels. However, before examining the nature of such integration, it will be useful to consider one type of mechanism by which an affectivity heuristic might be implemented.

Attractor Systems. Attractor systems provide a versatile form of representation that can be used to model the dynamics of a wide range of phenomena: from a ball rolling down a hill, and planets revolving around a star, to chaotic dynamics, including the growth of cities (Nicolis & Prigogine 1989; Morrison 1991; Dendrinos 1996). The generic flexibility that attractors empirically manifest in nature and society makes them useful as design components in agent simulation models as well.

The components of an artificial attractor can be briefly summarized. A hyperdomain binds a set of dimensions that are interrelated by potential interaction (non-decomposability). A hyperdomain associated with an entity or an environment may contain one or more attractors that can act upon associated entities. In a spatial context, the prototypical effect of such attraction might be to change the location of the entity within the hyperdomain.

An attractor must be defined in terms of at least one dimension, and may be defined on multiple dimensions, with the caveat that each dimension be part of the same hyperdomain. Agents must be defined in terms of: 1) the hyperdomains with which they have a composition relationship, and 2) attractors that influence them, and 3) the hyperdomains within which those attractors are defined. The effect of a given attractor may vary by agent type.

Consider a simple spatial framework. A small set of (minimally) ordinal, and typically orthogonal, dimensions is aggregated into a hyperdomain. These dimensions afford the possibility of movement, and thus introduce a "degree" of freedom for entities associated with the hyperdomain, although that degree may be uneven or idiosyncratic.

Points of attraction (repulsion) within a hyperdomain are defined upon one or more of the participating dimensions, and may manifest linear, cyclical, or fractal (or other nonlinear) forms of relations.[1] These attractors may have divergent effects based upon the entity type or the value region of the dimension, or both. Attractors may be defined on multiple participating dimensions, and there may be interaction effects between (among) those dimensions.

Integration of Agent Layers. An affectivity heuristic, implemented by proximity attractor, has the potential to integrate Muller's three agent layers. The first form of integration is achieved by implementing reactive and deliberative layers using a common mechanism. The mechanism is: 1) an attractor system, 2) coupled with simple response rules (e.g., reciprocity, the enemy of my enemy is my friend, etc.), and 3) supplemented by an inference mechanism. The situated attraction relations provide the first approximation for both the reactive and deliberative layers. However, at the reactive layer, there is little or no time for an inference mechanism to be invoked. At the deliberative layer, attractor relations provide a first approximation, an abduction to which situated inference can be applied. In particular, inconsistencies, conflict and areas of conceptual or plan incoherence can become the focus of agent deliberation.[2]

A second form of integration is achieved by implementing deliberative and interactive layers using the same model and mechanism. Inasmuch as agents occupy neighborhoods defined by spatial location, social networks, belief systems or other proximity relations, it is possible for attractor models to be shared in common. Consensual aspects of the attractor model can be represented once, with departures represented only for subgroups or individuals. This approach allows the explicit representation of communities of opinion and, at the same time, conserves computational resources.

Bounded Rationality. In a step toward making game theory models more realistic, Simon (1985) introduced the concept of bounded rationality. However, the utility of models that incorporate bounded rationality has yet to be established. As Rubinstein (1998, p. 3) recently observed, " . . . many of us feel that attempts to model bounded rationality have yet to find the right track." The model and mechanism described above are sufficiently flexible and expressive to contribute to the representation of bounded rationality in agents.

---

[1] For a mathematical description of attractor systems see Hurley (1982) and Alligood & Yorke (1989). For a qualitative analysis of their application to social systems, see Sallach (forthcoming).

[2] One practical advantage of this approach is a reduction in the computational costs of employing formal logic without foregoing the benefits.

They also have the potential to modify assumptions regarding the nature of bounded rationality. In Carley and Newell's view (1994), for example, social agents are best conceptualized by a *contraction* of the knowledge and capabilities available to an omnipotent agent. No longer omniscient with respect to the task environment, the boundedly rational agent knows less and has more limited processing capabilities. The authors describe emotions as degrading or limiting, rather than augmenting, human reasoning abilities (p. 225).

However, agents whose rationality is bounded are likely to have not only limited processing capabilities, but also positive compensations that accompany those limitations. It might be hypothesized that the effective integration of emotion with memory and strategic inference, as represented in the proposed model/mechanism, plays an important role in the activity of such a compensatory effect. Or, at least, it may contribute to the integration of agent responses in circumstances that require (permit) varying levels of inference.

References.

Alligood, Kathleen T., & James A. Yorke. 1989. Fractal basin boundaries and chaotic attractors. Pp. 41-55 in R.L. Devaney and L. Keen, eds., *Chaos and Fractals: The Mathematics behind Computer Graphics.* Providence: American Mathematical Society.

Brady, H.E. & P.M. Sniderman. 1985. Attitude attribution: A group basis for political reasoning. *American Political Science Review* 79:1061-1076.

Carley, Kathleen & Allen Newell. 1994. The nature of the social agent. *Journal of Mathematical Sociology* 19:221-262.

Dendrinos, D. S. 1996. Cities as spatial chaotic attractors. In L.D. Kiel & E. Elliott, eds., *Chaos Theory in the Social Sciences* (pp. 237-269). Ann Arbor: University of Michigan Press.

Hurley, M. 1982. Attractors: Persistence, and density of their basins. *Transactions of the American Mathematical Society 269* (January):247-271.

Morrison, Foster. 1991. *The Art of Modeling Dynamic Systems: Forecasting for Chaos, Randomness and Determinism.* New York: John Wiley.

Muller, Jorg P. 1996. *The Design of Intelligent Agents.* New York: Springer.

Nicolis, G., & Prigogine, I. 1989. *Exploring Complexity: An Introduction.* New York: W.H. Freeman.

Rubinstein, Ariel. 1998. *Modeling Bounded Rationality.* Cambridge, MA: MIT Press.

Sallach, David L. 1988. The structure of social mind: Empirical patterns of large-scale knowledge organization. Pp. 710-716 in the *Proceedings of the Tenth Annual Conference of the Cognitive Science Society*, Montreal.

_____. Forthcoming. Classical social processes: Attractor and computational models. *Journal of Mathematical Sociology.*

Simon, Herbert A. 1985. Human nature in politics: The dialogue of psychology with political science. *American Political Science Review* 79 (June):293-304.