

Modeling social relationship by mutual commitment, mutual control and mutual regulation

Alan H. Bond, Ph.D.,

California Institute of Technology, Mailstop 256-80,
1201, East California Boulevard, Pasadena, CA 91125

email: bond@cs.caltech.edu

webpage: <http://www.cs.caltech.edu/~bond>

Abstract:

We describe three possible levels and mechanisms for social relationship among agents. First, there is the level of negotiated commitment, by which agents by agreement mutually constrain their behavior to provide social coherence. Second, we describe how our agent model can be used to model mutual control among agents at different cognitive levels. This uses a parallel and modular agent architecture we have designed and implemented and which is based on blackboard ideas and inspired by the primate brain. Third, we discuss the concept of mutual regulation which operates at the level of resources.

1 Introduction

By social agent we mean an agent which represents other agents, which forms relationships with others by establishing joint commitments, and which enters into joint planning and joint action with others in its society (BG88). This social approach, it is believed, can result in better distributed systems and better agent architectures since it will build in automatically social effects which will inevitably be eventually required. We will need mechanisms for collaborative learning, for joint co-construction of action and for joint interpretation of experience.

We consider how a set of agents coheres as a system, and is able to function efficiently in a shared environment of knowledge and expectation. We seek to ground our ideas of multiagents in the concept of a surviving self-maintaining society. They need to maintain social relationships to assure survival of the group or system. This involves knowledge concerning other social agents and social relationships among them. The survival of the group, and the maintenance of social relationships, requires the maintenance and sharing of knowledge. In fact, of course the ability to share knowledge amplifies the abilities of the group and of the individuals of the group. Instead of detailed knowledge

of the knowledge of other social agents, a society maintains a common shared set of commitments.

Societies in nature tend to have a detailed, intimate, sharing of knowledge and a detailed set of shared commitments. Commitments include expectations as to what will occur, what others will tend to do, and so on. When such detailed sharing occurs, communication and shared action become very efficient. Social agents can predict each others actions and can thus greatly optimize their own resource use. Communication becomes efficient, allowing shared knowledge to be maintained at desirably high levels.

2 Mutual commitments

Social commitments can perhaps be divided into cultural and individual commitments. By cultural commitments we mean the many implicit concepts, assumptions and practices which are shared by all agents, and these can perhaps be ordered along an axis of negotiability, from the mythic to the everyday. In a computational model, their analog must be the assumptions built into the agent's mechanism by the designer and implementor.

We have proposed that individual commitments (Bon90) be modeled computationally by mutual agreement among a set of agents. Our intention was to model those working arrangements and organizations which can be explicitly represented. A formulation and computational model of the basic notion of commitment were given, which represented commitments as mutually agreed constraints on action, belief and world state. The constraints involved could be of various types, notably to follow a certain course of action, or to hold certain beliefs. A commitment concerns either acting in a certain way, conditional upon circumstances, or it can be a commitment to hold a certain belief. *Prospective* commitments concern the future, *retrospective* commitments concern the past. A retrospective commitment is a set of beliefs about past events. Commitments of course may be *conditional* on the state of the world, including the time. A

commitment concerning another agent's action or belief can be called an *expectation*. Our concept of commitment was that an agent is constructed to abide by its commitments; they could only be changed by renegotiation with the agents originally involved in creating the commitments.

Different computational mechanisms are required for the realization of these concepts. First, we need a negotiation mechanism by which agents can compute representations of constraints. We have ourselves only investigated logical negotiation, which we take to be joint proof conducted under conditions of complete truth and trust. Second, we need a mechanism for ensuring that commitments to action are carried out as agreed. And third, we need every action of the agent to be automatically constrained by any agreed commitments on belief and behavior. This latter is currently either extremely expensive or must be limited to certain tractable classes of constraint such as linear inequalities on numerical variables, as in constraint logic programming.

3 Mutual control

We implemented an architecture for a social agent as a hierarchy of modules which are executed in parallel, each module containing a store of logical descriptions and a set of rules which are executed in parallel (Bon99). Figure 1 depicts the architecture and its relation to the primate brain.

The action of this architecture is for each agent to continuously generate goals and from these to continuously select and instantiate a *joint plan*. We define a joint plan as a set of joint plan steps governed by temporal ordering and causality relations. One joint plan step contains descriptions of actions for each of a set of cooperating agents. This step is executed by perceiving the other cooperating agents and instantiating variables from the percept, and then instantiating, elaborating and executing the corresponding action for the subject agent.

The architecture forms a perception-action hierarchy, so that perception and action is described by expressions at several levels which condition one another. Thus action is conditioned by the current percept at the same level of abstraction, and perception is focussed by the requirements of action at the same level. The joint action is "voluntary" or "mixed initiative" in the sense that each agent is only controlled by its own goals, plans and knowledge.

This architecture is flexibly both goal-directed and environmentally situated. It is also quite appropriate for social interaction, since the other agents are perceived at each level and can directly influence the action of the subject agent. It allows agents to enter into stable mutually controlled behaviors where each is perceived to be carrying out

the requirements of the joint plan of the other. Further, this mutually controlled activity is hierarchically organized, in the sense that control actions fall into a hierarchy of abstraction, from easily altered details to major changes in policy.

We implemented two kinds of social behavior, one was affiliation in which agents maintained occasional face-to-face interactions which boosted affiliation measures, and the other was social spacing in which agents attempted to maintain socially appropriate spatial relationships characterized by proximity, displacement and mutual observability.

Figure 2 shows the instantaneous behavioral states of two interacting agents. Each is elaborating its selected joint plan conditional upon its perception of the other. We attempt in Figure 3 to depict two social agents as forming feedback loops. This is an attempt to depict mutual control. Each agent perceives both agents, which has some common and some private perception as input, and each agent executes its part of the joint action. These loops are social:

1. loop 1 maintains their current awareness and activation of their detailed physical actions
2. loop 2 maintains their current joint action or plan
3. loop 3 maintains their current joint goal or goals
4. loop 4 maintains the affiliative relationship between the set of two or more socially interacting agents.

Our notion of joint plan has some subtlety and indirectness, which is really necessitated by the distributed nature of agent interaction. Each participant has their own joint plan, which includes expectations of the actions of coparticipants. Each participant attempts to find and to carry out their "best" joint plan which satisfies their goals. There is no agreed shared plan as such. In constrained situations, it may be that the best joint plan of each participant is very similar to the best joint plan of coparticipants. Thus joint plans of individuals may be more or less engaged.

There are two standard examples: (1) the prostitute and the client, which coparticipate and cooperate, each with his or her own goals and joint plan. (2) adversarial coparticipation occurs with lawyers representing defendant and plaintiff. Since however there is always a residual conflict or disparity and residual shared benefits in all relationships, it is difficult to find cases of pure cooperation or even pure adversality.

The initiation (and termination) of joint action usually involves more disengagement between the joint plans of coparticipants. The grooming precludes observed in social monkeys are for example initially more unilateral. Initiation and termination usually involve protocols by which copartici-

pants navigate paths through a space of states of different degrees of mutual (dis)engagement.

In this view, social interaction is never unilateral. First, the other is always an imagined coparticipant. Second, even in the case of hardwired evolved behaviors, the behavior is intended for, only works with, and only makes sense, with a coparticipant, even though, in this case, there is no explicit representation of the other. It is not clear for example what representation, if any, of the mother a baby may have. There is for example biological evidence of tuning of the babies sensory systems during pregnancy, and immediately after birth, for example to the mother's odor and voice.

4 Mutual regulation

This mutually relating view allows us to connect our thinking with that of Myron Hofer (Hof84), (Hof87), who observed and described the mutually regulative nature of physiological mechanisms of mother and newborn, in the rat. These mechanisms maintain vital bodily functions such as warmth and nutrition, as well as heart rate, hormone levels and arousal levels. Thus the newborn's actions regulate the physiology of the mother and the mother's actions regulate the physiology of the newborn. Deviations from comfortable levels of some physiological parameter lead to actions which restore those parameters to comfortable levels. Such mechanisms are postulated by Pipp and Harmon (PH87) and by Kraemer (KESM91) as the basis of attachment. However, we are of the opinion that additional higher-level mechanisms are needed to give a satisfactory explanation of attachment behavior in primates, and certainly humans. The additional mechanisms involve constructing and maintaining higher-level representations in the child of the caretaker and in the caretaker of the child.

We can contemplate at least the possibility of computational mechanisms which would regulate resources of agents in such a subliminal way. This might correspond to runtime priority levels, swap space, disk allocations, bandwidth, and so on. In the case of human-computer interaction, it may also be possible to alter parameters of the human's environment. Whether this would have any use in system design remains to be seen.

Acknowledgements. This work was supported by the National Science Foundation, Computation and Social Systems Program, Research grant IIS-9812714, and by the Caltech Neuromorphic Engineering Research Center. The author would like to thank Professors Pietro Perona, Mani Chandy and Steven Mayo for their support and encouragement.

References

- Alan H. Bond and Les Gasser. *Readings in Distributed Artificial Intelligence*. Morgan Kaufmann Publishers, San Mateo, CA, 1988.
- Alan H. Bond. Commitment: A Computational Model for Organizations of Cooperating Intelligent Agents. In *Proceedings of the 1990 Conference on Office Information Systems*, pages 21–30, Cambridge, MA, April 1990.
- Alan H. Bond. Describing Behavioral States using a System Model of the Primate Brain. *American Journal of Primatology*, 49(4):315–338, 1999.
- Myron A. Hofer. Relationships as regulators: A psychobiologic perspective on bereavement. *Psychosomatic Medicine*, 46:183–197, 1984.
- Myron A. Hofer. Early social relationships: a psychobiologist's view. *Child Development*, 58:633–647, 1987.
- G. W. Kraemer, M. H. Ebert, D. E. Schmidt, and W. T. McKinney. Strangers in a strange land: a psychobiological study of infant monkeys before and after separation from real or inanimate mothers. *Child Development*, 62:548–566, 1991.
- Sandra Pipp and Robert J. Harmon. Attachment as regulation: a commentary. *Child Development*, 58:648–652, 1987.

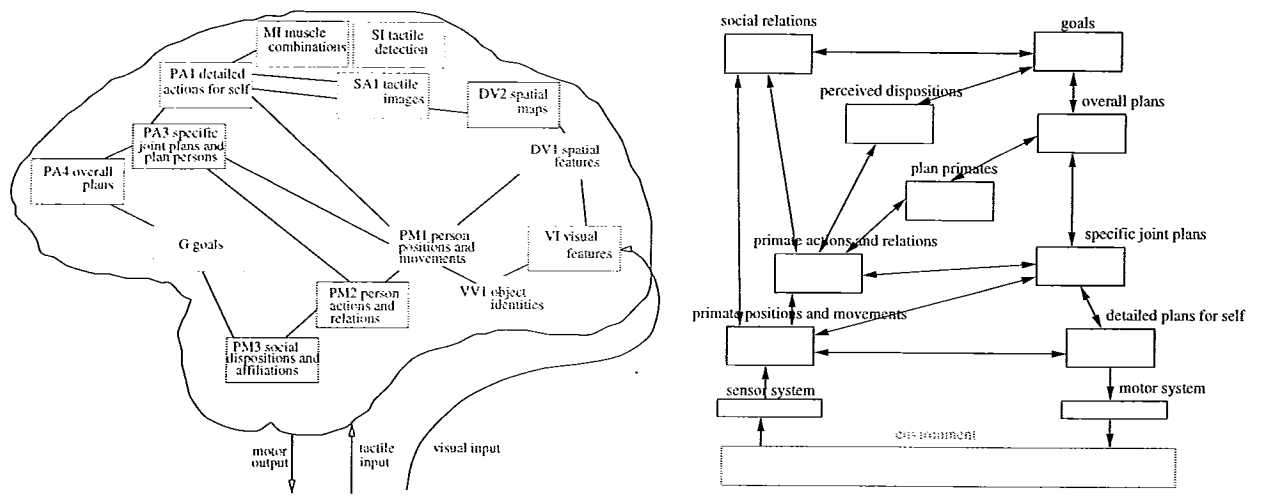


Figure 1. Modules from neural areas of the primate neocortex, and our initial system model

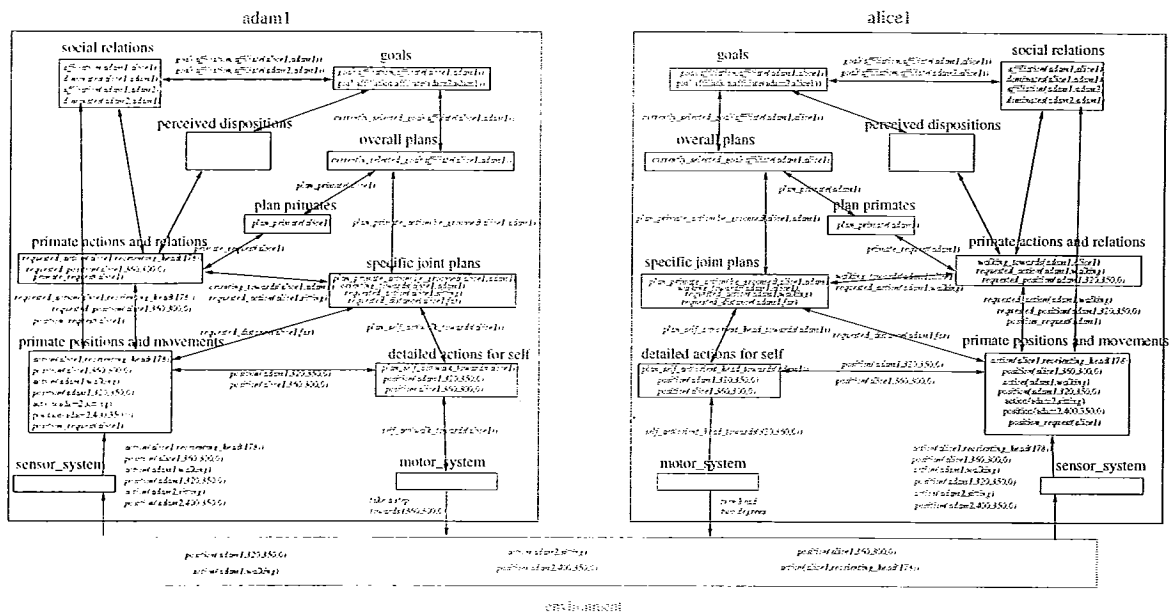


Figure 2. Instantaneous behavioral states of two interacting agents

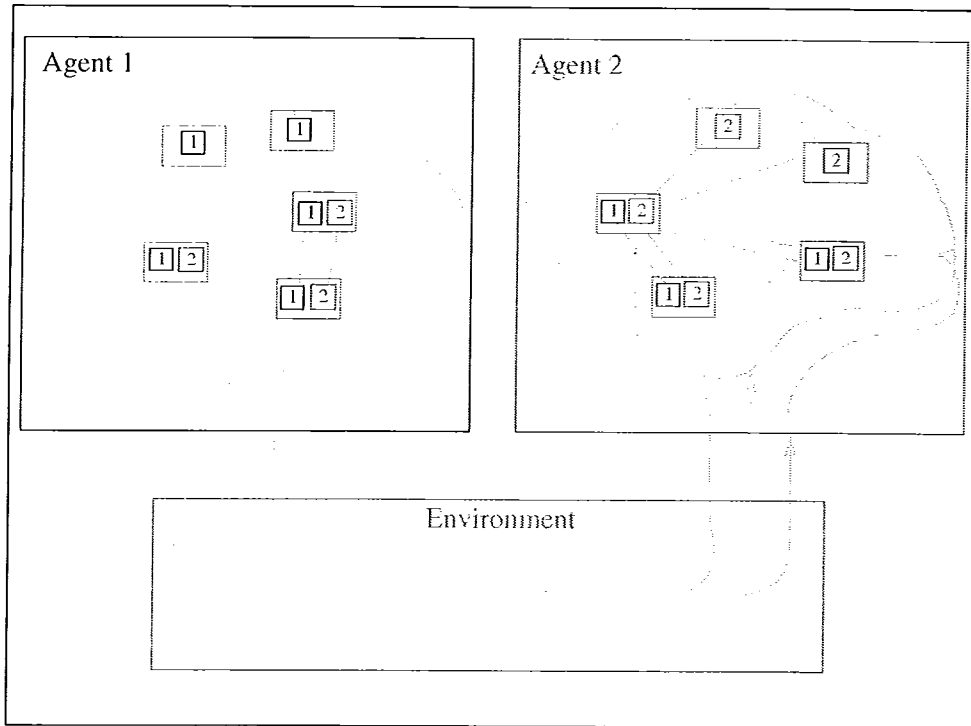


Figure 3: Mutual control