A Conceptual Analysis of Commitments in Multiagent Systems

Munindar P. Singh*
Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206, USA

singh@ncsu.edu

May 16, 1996

Abstract

The notion of commitment is central to understanding agents and multiagent systems. At least two kinds of commitment can be identified in the AI literature—the internal or psychological and the external or social. While these notions must not be conflated with each other, they are not entirely unrelated. We review the historical development of commitments in AI and distributed computing, and show the various roles they might play in multiagent systems. We discuss the key interrelationships among these concepts, and study their implementational aspects, both in traditional individual agents, and in agents that are recursively created as systems of other agents. We close with a discussion of the key research challenges.

*This is a considerably revised version of a DFKI Technical Memo. Parts of it were presented at a panel discussion at the 1991 AAAI Fall Symposium on Knowledge and Action at the Social and Organizational Levels. An invited presentation in a panel on *Agents and Commitments* at the 1995 International Conference on Multiagent Systems (ICMAS) is also incorporated.

Contents

1 Introduction

The notion of commitment is central to much recent work in AI. Commitment, however, comes in two main flavors. One is the familiar notion from ethics and distributed computing where an agent makes a commitment to others to do or not do certain actions. The agent is liable for not acting up on them. The other sense of commitment involves an agent by himself. The agent is committed to his intentions or beliefs, but is not liable (to anyone else) [Bratman, 1987, ch. 2] [Harman, 1986, p. 94] [Cohen & Levesque, 1990, p. 217] [Singh, 1991b]. This is related to epistemic entrenchment [Gärdenfors, 1988]. We call the first kind of commitment *S-commitment* (for social) and the second kind *P-commitment* (for psychological).

Historically, commitments were of the psychological, intra-agent variety. They were procedurally—and unwittingly—encoded in planning algorithms, which led to suboptimal solutions. This generated a research interest in *least-commitment* planning [Sacerdoti, 1977], wherein commitments were actively avoided. But, as the reasoning limitations of agents were recognized, commitments became recognized as a means to control search and obtain satisficing, instead of optimal, solutions. Thus commitment became a "feature"! Commitments have since been extended to distributed settings—we shall review their ramifications in detail below. Although we concentrate on the main flavors of commitments, other dichotomies exist, e.g., epistemic versus conative, and commitments to actions versus commitments to policies.

Before we get into the details, it is instructive to review the commitment as it is implemented in distributed computing. This is one of its mundane variations, but will give us a base-line for comparing more sophisticated variants, because it is a simple situation with many of the interesting ingredients. Mutual commit protocols ensure that a number of distributed processes agree on some important action. An important application is whether a given distributed database transaction should commit at all sites, which is essential to preserving the integrity of distributed data [Gray & Reuter, 1993]. The most famous of these protocols is two-phase commit (2PC). In 2PC, there is a voting phase in which each process votes Yes or No. The votes are collected by a coordinator process. If all processes voted Yes, the coordinator announces a Yes decision, otherwise a No decision.

First, the above phenomenon really is commitment, since the processes guarantee that they will behave as the coordinator decided. However, this commitment is a one-shot affair. The processes agree on whatever they are doing, do it, and move on. The commitment they exhibit is irrevocable, but since the interaction is short-lived, that cannot cause significant problems. Second, the commitment is represented trivially. There is typically a flag set inside each process. Arguably, this flag represents a P-commitment on part of the process. The S-commitment aspects are provided externally, through an agreement among the various programmers. This agreement ensures that processes will respect the decision of the coordinator, and once they

commit to a decision, they will not change their mind. Thus each site knows that the other sites will behave as it expects.

However, the above hardwired commitment is at a low level of abstraction. It provides no flexibility in terms of behavior. Typically, even the physical representation is fixed. Not only AI, but also traditional applications, e.g., collaborative computing and workflows, demand greater organizational flexibility [Elmagarmid, 1992; Papazoglou *et al.*, 1992].

A higher level of abstraction must involve agents. Agents are systems that can be understood through psychological and social abstractions with the *intentional stance* [McCarthy, 1979] or equivalently at the *knowledge level* [Newell, 1982]. This view, which justifies our present terminology of P-commitments and S-commitments, is crucial to the AI enterprise. We submit that P-commitments and S-commitments will provide two of the most important abstractions in characterizing, understanding, analyzing, and designing multiagent systems. P-commitments have been shown to underlie traditional notions such as belief and intentions (we elaborate this role further here). S-commitments play a similarly important role in coordinating and structuring multiagent systems, and in achieving coherence in their actions.

The challenges we address in this paper are to

- lay out the conceptual underpinnings of P-commitment and S-commitment in resource-bounded multiagent systems
- map out the terrain of the above concepts in terms of their interrelationships in different situations
- relate the above concepts to the key features in the design and architecture of agents and multiagent systems.

Section 2 discusses the two notions of commitment in some detail. Section 3 relates them to the structure of multiagent systems. Section 4 discusses their implementational import. Section 5 concludes with a review of the pertinent literature.

2 Commitments

P-commitments arise within agents. S-commitments arise when agents interact with other agents. Just as P-commitments are needed to understand individual agents, S-commitments are needed to understand multiagent systems. Traditional formal theories give preeminence to P-commitment and do not recognize the complexities of S-commitment possible in real systems—they allow S-commitments implicitly but only in systems which are internally homogeneous. In this section, we shall study the two concepts as first-class concepts, and outline the various properties they have and roles they play. From this analysis, we will be better placed to understand their

interrelationships. We use the term *commiter* to refer to the agent that makes a commitment, and the term *commitee* (not "committee") to refer to the agent who receives the commitment. There is no commitee for P-commitments.

Our motivation is both descriptive and prescriptive. First, commitments of both kinds are held by the intelligent systems that we encounter in real-life: humans, corporations, etc. Second, commitments of both kinds can be motivated on grounds of individual and group rationality for limited systems. We examine the relationship between each kind of commitments and issues such as individual rationality, group rationality, autonomy, representation, and coordination. The issue of interfaces applies only to S-commitments.

2.1 Psychological Commitments

We mentioned how P-commitments came to be appreciated in AI. One of their major applications is in understanding intentions. Intentions for future actions are an important concept. Perhaps their salient property is that they involve a commitment on the part of agents. Intentions are taken to be causes of actions, and can involve actions in the future. This requires some persistence or commitment. This has been supported by several researchers [Bratman, 1987, ch. 2], [Harman, 1986, p. 94], [Pollack, 1992], and [Cohen & Levesque, 1990, p. 217]. An agent who has an intention is in some way committed to it—not only does he intend to achieve the relevant condition right now, but would also intend to achieve it later, even as the circumstances changed, perhaps for the worse.

Agents can also be committed toward their beliefs. This commitment corresponds to doxastic or epistemic entrenchment [Gärdenfors, 1988, ch. 3], which means that the more committed an agent is, the more he will resist reconsideration, even as he receives new evidence. For example, if you firmly believe that all birds can fly, you might accept that penguins aren't birds, rather than that some birds cannot fly.

This suggests the unifying role of P-commitments, whether for intentions or beliefs. P-commitments constrain the deliberation of an agent about his cognitive state. Thus, they provide an important component of the architecture of agents viewed at the knowledge level.

2.1.1 P-commitments and individual rationality

P-commitment entails that the agent continue with a belief that might no longer be justified and an intention that might no longer be feasible or serve the agent's ultimate goals. Thus there is a certain amount of irrationality built into the very concept of commitment. But, for limited agents, commitments can be reconciled with rationality. Limited agents lack the resources to reason at every moment from "first principles"—this problem is especially acute for agents who must act in real-

time and who live in a complex environment. P-commitments help the agent consider only a few issues, which he can hope to resolve in the limited time he has available. Well-designed agents have the relevant commitments so that they can succeed despite ignoring a number of aspects they might otherwise reason about.

2.1.2 P-commitments and group rationality

P-commitments have only an indirect effect on group rationality. In cooperative systems, this is essentially because P-commitments add to the coordination among the agents. Agents with P-commitments project a stable image and are much easier to predict the behavior of by other agents. Thus the limitations of the agents, which increase the extent of their P-commitments, also make them easier to predict. However, in non-cooperative settings, agents could attempt to exploit the predictability of each other to maximize their individual gains. This leads to the classical prisoners' dilemma, which we revisit below in the context of S-commitments.

2.1.3 P-commitments and autonomy

Every commitment translates to a reduction in the choices are available to the agent. However, in the case of P-commitments, this reduction is entirely voluntary, and as we discuss below, can be unilaterally revoked. Thus, we claim there is no loss in autonomy.

2.1.4 P-commitments and representation

In typical systems, P-commitments are explicitly symbolically represented as lists of beliefs or intentions. The proper commitment aspect is in the mode of reasoning of the agent. However, in general, P-commitments can arise both with and without explicit representation, depending on the kind of agent.

2.1.5 P-commitments and coordination

This is similar to the discussion on group rationality. P-commitments by themselves do not assist in coordination, but in conjunction with other system properties (e.g., cooperation), they can enhance coordination.

2.2 Social Commitments

This kind of commitment arises in distributed AI, e.g., when the agent promises (to another) to perform a certain action. The S-commitment of an agent (to another) in a multiagent system to achieve a certain goal is essential to coordination. For example, if one agent promises another to be at a rendezvous spot, then the other can take

this for granted and be there himself; or if one agent promises to lift one end of a piano, the other can lift the other end, and jointly they can lift the piano, thereby achieving something they could not have achieved singly. It is important to note that commitments in classical distributed computing are irrevocable—this makes much of this research inapplicable to AI, where such rigidity is undesirable.

S-commitments are at least in practice an important concept for AI. It would be impossible to design or understand sufficiently complex systems of autonomous agents without invoking the concept of S-commitment. The agents' S-commitments give us a powerful abstraction with which to view their interactions. Intuitively, just as traditionally psychological concepts are needed for agents, social concepts are needed for systems of agents. And, S-commitment is a social concept *par excellence*.

2.2.1 S-commitments as interfaces

S-commitments are a way of specifying interfaces among intelligent agents in multiagent systems. The most interesting interactions among agents involve communications. Communications can be described fruitfully using social commitments. For example, promises ordinarily bring into effect a social commitment by the speaker to the hearer; directives presuppose a S-commitment by the hearer to do as told; assertives S-commit the speaker to the statement expressed; permissions make the speaker S-committed to allowing the relevant condition to hold; prohibitions presuppose a S-commitment by the hearer to preventing the relevant condition from holding. Notice that promises and assertives bring about the relevant S-commitments, even if the speaker is *insincere*! The speaker is subject to censure for having lied to the listener.

Indeed, S-commitments differentiate communication from other interactions, such as resource conflicts. Although an agent might obtain some information from another agent because of their conflicts, these interactions differ from the proper communicative ones in lacking the properties discussed above. Interaction protocols, e.g., for various kinds of negotiation, between different agents may be defined so that the interacting agents have the relevant S-commitments. This generalizes the idea of [Singh, 1994] that participating agents have the requisite intentions and know-how.

A practical consequence is that one can design agents independently of each other and just ensure that their S-commitments mesh in properly when they are combined.

2.2.2 S-commitments and individual rationality

Surprisingly perhaps, S-commitments have a natural connection with individual rationality. They assist in the setting up of various negotiations, which enable bargaining between agents. Bargaining can never be about what P-commitments an agent will adopt, but must be about his S-commitments. If agent x agrees to do something that another agent y intends he do, x might obtain some payment from y. If x just

intended to do what y desires, but does not S-commit to it, x would have no grounds for bargaining.

In addition, S-commitments can make it easier for the group as a whole to achieve its ends, and this might have some positive utility for a given agent, as an individual.

2.2.3 S-commitments and group rationality

The role of S-commitments in group rationality is akin to that of P-commitments in individual rationality. Even when they are not in the immediate (i.e., local) interest of an agent, S-commitments might be a good idea from the systems point of view. When agents in a system make and keep S-commitments, the system can perform better than it might if they each worked selfishly. This too depends on specific systems, but there are numerous real-life systems where S-commitments alone make it possible for any kind of success. For example, the famous prisoners' dilemma paradox arises when the individually preferred actions of two prisoners (namely to tell on the other) lead to a worse pay-off for each than the individually less preferred actions of protecting the other. The latter actions can be rationally performed only if the agents have S-commitments to not aid the police.

2.2.4 S-commitments and autonomy

Just as P-commitments, S-commitments reduce an agent's choices. However, because of their interagent aspects, S-commitments can be unilaterally revoked in most cases. This means that they are a genuine restriction of the commiter's autonomy. This is an instance of what we believe is a fundamental trade-off between coordination and autonomy. Any kind of coordination entails a loss of freedom in the narrow sense. Of course, an agent could decide sufficiently in advance not to make any S-commitments and then won't have to worry about losing any choices on their account.

2.2.5 S-commitments and representation

S-commitments may be implicit. That is, they need not be explicitly symbolically represented by the agents, but could instead be derived from their social architectures. Of course, they *could* be symbolically represented by smart introspective agents. Often the relevant agents would be treated only indexically, i.e., relative to the agent's own situation. For example, at a stop sign at a street crossing, the other agent would be "the driver to my right," rather than Bill, the councilman. The S-commitment itself to let him go (and the S-commitment on his part to get out of my way) are both a matter of training, here, of learning how to drive. S-commitments may also be generated by the *social roles* of agents. For example, someone, merely by being a policeman in uniform, is S-committed to chase after a criminal. His not doing so would be a dereliction of duty.

2.2.6 S-commitments and coordination

S-commitments enable a simple definition of coordination: the involved agents merely have to have the appropriate S-commitments to one another. Two agents are jointly committed to achieving a condition, p, iff each of them is S-committed to the other to achieve a (possibly different) condition such that the achievement of those two conditions would entail p. Thus each agent could act on the assumption that the other would do his share.

2.3 Adopting and Revoking Commitments

Whereas, P-commitments may be established autonomously by the given agent S-commitments are established by participation in certain social situations or by making the appropriate speech acts. Castelfranchi proposes that S-commitments are established in the presence of a *witness*, who as it were officiates at the event [Castelfranchi, 1993]. In addition, in many cases when several agents are simultaneously entering into related S-commitments with each other, it is appropriate to include a cancelation clause, which limits the liability of the participants, and determines how the deal is to be broken if ever.

Ideally, both kinds of commitments are dropped when they complete successfully, e.g., when the intended or promised condition is achieved. However, we must consider when they are prematurely dropped or revoked.

In principle, both kinds of commitment can be revoked at will. However, in the case of P-commitments, the agent is truly unconstrained as to when he drops them. In the case of S-commitments, however, the agent is constrained. This could be because the initial agreement through which the agent entered into the given S-commitment restricts his freedom, or it could because applicable rules or conventions do so. In some cases, the commitee would be judge of whether the S-commitment was successfully fulfilled; more often, it would be an external authority, e.g., the witness.

Prescriptively, we would like to prevent unilateral revocation of S-commitments (but not always—e.g., someone who foolishly agrees to rob a bank might be forgiven for not showing up). A plausible requirement is that the commiter obtain an acknowledgment from the commitee.

2.4 Relating S-Commitments and P-Commitments

It is obvious that P-commitments and S-commitments are distinct concepts.

- An agent may be P-committed to a belief or an intention and yet not be S-committed to it. For example, an agent might have some secret beliefs that he has not revealed to anyone.

- An agent might be officially S-committed to an intention or belief, but not be P-committed to it. For example, a bad cop might not intend to catch a robber friend of his. He may willingly pay the penalty or face censure.

- An agent, e.g., Robinson Crusoe, can have P-commitments, by himself, e.g., to trap a goat. But he needs another agent, e.g., Man Friday, to interact with to have S-commitments, e.g., to have dinner at sundown.

Despite the above, however, felicitously having a S-commitment presupposes having the corresponding P-commitment. At the least, an agent who is S-committed to achieving something should intend to achieve it. If agents weren't P-committed to their S-commitments, they would be more likely to fail to satisfy them. Although, they may be penalized for this, the advantages of S-commitments enumerated in section 2.2 would be lost.

P-commitments do not *generate* S-commitments, even if they are a reason why a rational selfish agent enters into S-commitments. S-commitments can exist only because of the social relationship between the concerned agents, even if it is entered into consciously and deliberately by them. Consequently, S-commitments are fundamentally more complex than P-commitments.

It is clear that S-commitments cannot be trivially reduced to P-commitments. Some researchers, however, have proposed formal definitions that assume that S-commitments can be reduced to statements involving mutual beliefs among the participating agents. We review these suggestions in section 3. In the opposite direction, it has been proposed to P-commitments reduce to S-commitments to oneself [Shoham, 1993]. This reduction defines the commitment or obligation of x to x the same way it defines the obligation of x to y. This misses the crucial intuition of the representation of self, called the *essential indexical* in [Perry, 1979]. Perry shows how an agent's actions depend on whether he knows the "other" party is himself. Thus the reduction can succeed only if an additional primitive of *self* is added.

3 Commitments and Structure

Commitments have an intimate relationship with the structure of multiagent systems. Although it is obvious that multiagent systems are structured, relatively few theories can actually accommodate it. We have identified two principles in [Singh, 1991a] that are relevant here.

- **Heterogeneity.** Systems are composed of a diverse mix of agents and other systems; the members may have different knowledge, know-how, intentions, etc. In some cases, they may not be aware of each others' existence, or know each others' names.

- **Monolithicity.** Despite this, systems can be viewed as if they were single agents. They provide a single locus of actions and can have beliefs, intentions, etc. They can enter into S-commitments. Groups such as teams, armies, nations and corporations may all be (and are in fact) profitably treated as being single, though potentially complex, agents.

We developed a rigorous theory of abstractions for multiagent systems in which the structure is indirectly captured through constraints on the communications among member agents [Singh, 1994]. Those results can be refined through the use of S-commitments, related to the proposal on communications in section 2.2. Thus the structure of a system might be captured directly through the S-commitments among its members. In fact, behaviorally identical systems could have vastly differing structures. For example, one structure may include a central authority and another might not, with obviously different applicability.

The idea of structured and recursively composed systems or agents combines with the notion of commitment to yield a powerful means of specifying properties of multiagent systems.

1. Systems may have P-commitments. These are cashed out in terms of the P-commitments and S-commitments of their members.

2. Systems may have S-commitments. These are reflected in their P-commitments and their internal structure in a way that is roughly analogous to, but more complex than, the case of individuals.

A useful paradigm is of S-committing an agent to a system that contains him. This enables capturing the common social situations, where one often speaks of abstract agents that other agents are committed to; these abstract agents include teams, nations, and other such entities.

Traditional construals of group action require the agents to have mutual beliefs about the joint action [Levesque *et al.*, 1990; Grosz & Sidner, 1990]. Mutual beliefs arise when each of a set of agents believes something, and believes that each of the other does so, and so on *ad infinitum*. This has a number of shortcomings.

- It assumes that the agents know of each other, and keep each other perfectly well-informed: thus it fails for systems whose membership is large and changing.

- It assumes that the system is homogeneous and symmetric with respect to the agents: each contributes the same in terms of what they know of the joint action

- It is fragile in that joint action becomes impossible whenever the mutual beliefs break down, which can happen whenever an agent gets the slightest doubt. More importantly, mutual beliefs are impossible to attain through communication over unreliable channels. They typically arise when they are designed in.

The use of mutual belief appears to be a technique to model S-commitments using P-commitments. We submit it is more appropriate to model the social concept directly. This is because mutual belief is not really interpreted in a manner similar to belief, and does not reduce the conceptual complexity of the descriptions.

4 Implementational Aspects

How can theoretical constructs like commitments be made to have some impact on the computations of the systems to which they apply. We are familiar with notions similar to P-commitments for individual agents, but the remaining possibilities are less clear. These are the implementation of P-commitments for systems (i.e., structured agents) and S-commitments for individuals and systems.

Even for the first case, there are some variations possible. P-commitments could be captured declaratively, i.e., symbolically, or procedurally. Often, the base concepts of intentions or beliefs are implemented declaratively, but the P-commitment itself lies in the processing of the base concepts. It would certainly help to make P-commitments themselves be declarative, for the usual reasons of making the implementation flexible with regard to how it reasoned and how the P-commitments were realized.

P-commitments for multiagent systems are more tricky, since they involve some coordination among the member agents. Depending on the structure of the system, this could be trivial or complex. In the latter case, P-commitments at one level could translate to S-commitments among agents of the lower level. For example, in an army the intentions of the army as a whole are given by the intentions of its general. By contrast, in a team, each member might make a small contribution.

S-commitments among individuals must rely upon some conventions about how they are processed. There is clearly a representation of the specific S-commitment by the commiter and commitee, but much of the work is in the external conventions. For example, in 2PC, all the processes can assume that the commitments are irrevocable.

S-commitments among multiagent systems pose no special problems. They still require representations by designated members of the systems, just as in the case of P-commitments. In addition, they require the conventions that S-commitments among individuals require, plus a notion of a designated member who has responsibility for the requisite actions.

The problems of joint actions by heterogeneous systems can be addressed. The members do not need to know the names of who else is involved, but if they have a representation for the system as an abstract agent, they can direct their S-commitments to it. This would apply very naturally for a large company, or an army, or a nation.

5 Comparisons and Conclusions

We now review some of the relevant literature. We referred to [Castelfranchi, 1993; Levesque *et al.*, 1990; Grosz & Sidner, 1990; Shoham, 1993] in the body of the paper. Gasser reviews some of the sociological issues underlying multiagent systems [Gasser, 1991]. His notion of the multiple simultaneous roles played by social agents inspired part of our discussion above. Castelfranchi studies concepts similar to those here [Castelfranchi, 1993]. However, he does not discuss the rationality and computational aspects of commitments. Also, he distinguishes a notion of collective commitment, which is subsumed by our concept of S-commitment through the orthogonal representation of the structure of multiagent systems. Tuomela develops an interesting theory of joint action and intention that bears similarities to collective commitments [Tuomela, 1991]. Several researchers are building systems that embody versions of S-commitment and exhibit the usefulness of this program of research [von Martial, 1992; Wittig, 1992].

We have sought to present the unifying principles behind two major notions of commitment for single-agent and multiagent systems. We have shown how these notions relate to various aspects of agent and system architecture, and the roles they might play in the science of multiagent systems.

Acknowledgements

This paper has been several years in the making. I have benefited from discussions with several people over the years, most notably Nicholas Asher, Cristiano Castelfranchi, Les Gasser, Michael Georgeff, and Michael Huhns. This work has been supported by a number of agencies, including the NSF, DFKI, MCC, and NCSU. I alone am responsible for the content, however.

References

[Bratman, 1987] Bratman, Michael E.; 1987. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA.

[Castelfranchi, 1993] Castelfranchi, Cristiano; 1993. Commitments: From individual intentions to groups and organizations. In *Proceedings of the AAAI-93 Workshop on AI and Theories of Groups and Organizations: Conceptual and Empirical Research.*

[Cohen & Levesque, 1990] Cohen, Philip R. and Levesque, Hector J.; 1990. Intention is choice with commitment. *Artificial Intelligence* 42:213–261.

[Elmagarmid, 1992] Elmagarmid, Ahmed K., editor. *Database Transaction Models for Advanced Applications.* Morgan Kaufmann.

[Gärdenfors, 1988] Gärdenfors, Peter; 1988. *Knowledge in Flux.* MIT Press, Cambridge, MA.

[Gasser, 1991] Gasser, Les; 1991. Social conceptions of knowledge and action: DAI foundations and open systems semantics. *Artificial Intelligence* 47:107–138.

[Gray & Reuter, 1993] Gray, Jim and Reuter, Andreas; 1993. *Transaction Processing: Concepts and Techniques.* Morgan Kaufmann.

[Grosz & Sidner, 1990] Grosz, Barbara and Sidner, Candace; 1990. Plans for discourse. In Cohen, P.; Morgan, J.; and Pollack, M., editors, *SDF Benchmark Series: Intentions in Communication.* MIT Press, Cambridge, MA.

[Harman, 1986] Harman, Gilbert; 1986. *Change in View.* MIT Press, Cambridge, MA.

[Levesque *et al.*, 1990] Levesque, H. J.; Cohen, P. R.; and Nunes, J. T.; 1990. On acting together. In *AAAI.*

[McCarthy, 1979] McCarthy, John; 1979. Ascribing mental qualities to machines. In Ringle, Martin, editor, *Philosophical Perspectives in Artificial Intelligence.* Harvester Press. Page nos. from a revised version, issued as a report in 1987.

[Newell, 1982] Newell, Allen; 1982. The knowledge level. *Artificial Intelligence* 18(1):87–127.

[Papazoglou *et al.*, 1992] Papazoglou, Mike P.; Laufmann, Steven C.; and Sellis, Timothy K.; 1992. An organizational framework for cooperating intelligent information systems. *International Journal on Intelligent and Cooperative Information Systems* 1(1):169–202.

[Perry, 1979] Perry, John; 1979. The problem of the essential indexical. *Nous* 13:3–21.

[Pollack, 1992] Pollack, Martha E.; 1992. The uses of plans. *Artificial Intelligence.* Computers and Thought Award Lecture.

[Sacerdoti, 1977] Sacerdoti, Earl; 1977. *The Structure of Plans and Behavior.* Elsevier Science Publishers.

[Shoham, 1993] Shoham, Yoav; 1993. Agent-oriented programming. *Artificial Intelligence* 60:51–92.

[Singh, 1991a] Singh, Munindar P.; 1991a. Group ability and structure. In Demazeau, Y. and Müller, J.-P., editors, *Decentralized Artificial Intelligence, Volume 2*. Elsevier Science Publishers B.V. / North-Holland, Amsterdam, Holland. 127–145.

[Singh, 1991b] Singh, Munindar P.; 1991b. Intentions, commitments and rationality. In *13th Annual Conference of the Cognitive Science Society.*

[Singh, 1994] Singh, Munindar P.; 1994. *Multiagent Systems: A Theoretical Framework for Intentions, Know-How, and Communications.* Springer Verlag, Heidelberg, Germany.

[Tuomela, 1991] Tuomela, Raimo; 1991. We will do it. an analysis of group-intentions. *Philosophy and Phenomenological Research* LI(2):249–277.

[von Martial, 1992] von Martial, Frank; 1992. *Coordinating Plans of Autonomous Agents.* Springer Verlag, Berlin, Germany.

[Wittig, 1992] Wittig, Thies, editor. *ARCHON: An Architecture for Multi-agent Systems.* Ellis Horwood Limited, West Sussex, UK.