Coordination Techniques for Distributed Artificial Intelligence

N. R. Jennings
Dept. of Electronic Engineering
Queen Mary and Westfield College
University of London
Mile End Rd.
London E1 4NS
UK
`N.R.Jennings@qmw.ac.uk`

Abstract

Coordination, the process by which an agent reasons about its local actions and the (anticipated) actions of others to try and ensure the community acts in a coherent manner, is perhaps the key problem of the discipline of Distributed Artificial Intelligence (DAI). In order to make advances it is important that the theories and principles which guide this central activity are uncovered and analysed in a systematic and rigourous manner. To this end, this paper models agent communities using a distributed goal search formalism, and argues that *commitments* (pledges to undertake a specific course of action) and *conventions* (means of monitoring commitments in changing circumstances) are the foundation of coordination in all DAI systems.

1. The Coordination Problem

Participation in any social situation should be both simultaneously constraining, in that agents must make a contribution to it, and yet enriching, in that participation provides resources and opportunities which would otherwise be unavailable (Gerson, 1976). Coordination, the process by which an agent reasons about its local actions and the (anticipated) actions of others to try and ensure the community acts in a coherent manner, is the key to achieving this objective. Without coordination the benefits of decentralised problem solving vanish and the community may quickly degenerate into a collection of chaotic, incohesive individuals. In more detail, the objectives of the coordination process are to ensure: that all necessary portions of the overall problem are included in the activities of at least one agent, that agents interact in a manner which permits their activities to be developed and integrated into an overall solution, that team members act in a purposeful and consistent manner, and that all of these objectives are achievable within the available computational and resource limitations (Lesser and Corkill, 1987). Specific examples of coordination activities include supplying timely information to needy agents, ensuring the actions of multiple actors are synchronised and avoiding redundant problem solving.

There are three main reasons why the actions of multiple agents need to be coordinated:

• because there are dependencies between agents' actions

Interdependence occurs when goals undertaken by individual agents are related - either because local decisions made by one agent have an impact on the decisions of other community members (eg when building a house, decisions about the size and location of rooms impacts upon the wiring and plumbing) or because of the possibility of harmful interactions amongst agents (eg two mobile robots may attempt to pass through a narrow exit simultaneously, resulting in a collision, damage to the robots and blockage of the exit).

• because there is a need to meet global constraints

Global constraints exist when the solution being developed by a group of agents must satisfy certain conditions if it is to be deemed successful. For instance, a house building team may have a budget of £250,000, a distributed monitoring system may have to react to critical events within 30 seconds and a distributed air traffic control system may have to control the planes with a fixed communication bandwidth. If individual agents acted in isolation and merely tried to optimise their local performance, then such overarching constraints are unlikely to be satisfied. Only through coordinated action will acceptable solutions be developed.

• because no one individual has sufficient competence, resources or information to solve the entire problem

Many problems cannot be solved by individuals working in isolation because they do not possess the necessary expertise, resources or information. Relevant examples include the tasks of lifting a heavy object, driving in a convoy and playing a symphony. It may be impractical or undesirable to permanently synthesize the necessary components into a single entity because of historical, political, physical or social constraints, therefore temporary alliances through cooperative problem solving may be the only way to proceed. Differing expertise may need to be combined to produce a result outside of the scope of any of the individual constituents (eg in medical diagnosis, knowledge about heart disease, blood disorders and respiratory problems may need to be combined to diagnose a patient's illness). Different agents may have different resources (eg processing power, memory and communications) which all need to be harnessed to solve a complex problem. Finally, different agents may have different information or viewpoints of a problem (eg in concurrent engineering systems, the same product may be viewed from a design, manufacturing and marketing perspective).

Even when individuals can work independently, meaning coordination is not essential, information discovered by one agent can be of sufficient use to another that the two agents can solve the problem more than twice as fast. For example, when searching for a lost object in a large area it is often better, though not essential, to do so as a team. Analysis of this "combinatorial implosion" phenomena (Kornfield and Hewitt, 1981) has resulted in the postulation that cooperative search, when sufficiently large, can display universal characteristics which are independent of the nature of either the individual processes or the particular domain being tackled (Clearwater *et al.*, 1991).

If all the agents in the system could have complete knowledge of the goals, actions and interactions of their fellow community members and could also have infinite processing power, it would be possible to know exactly what each agent was doing at present and what it is intending to do in the future. In such instances, it would be possible to avoid conflicting and redundant efforts and systems could be perfectly coordinated (Malone, 1987). However such complete knowledge is infeasible, in any community of reasonable complexity, because bandwidth limitations make it impossible for agents to be constantly informed of all developments. Even in modestly sized communities, a complete analysis to determine the detailed activities of each agent is impractical - the computation and communication costs of determining the optimal set and allocation of activities far outweighs the improvement in problem solving performance (Corkill and Lesser, 1986).

As all community members cannot have a complete and accurate perspective of the overall system, the next easiest way of ensuring coherent behaviour is to have one agent with a wider picture. This global controller could then direct the activities of the others, assign agents to tasks and focus problem solving to ensure coherent behaviour. However such an approach is often impractical in realistic applications because even keeping one agent informed of all the actions in the community would swamp the available bandwidth. Also the controller would become a severe communication bottleneck and would render the remaining components unusable if it failed.

To produce systems without bottlenecks and which exhibit graceful degradation of performance, most DAI research has concentrated on developing communities in which both control and data are distributed. Distributed control means that individuals have a degree of autonomy in generating new actions and in deciding which tasks to do next. When designing such systems it is important to ensure that agents spend the bulk of their time engaged on solving the domain level problems for which they were built, rather than in communication and coordination activities. To this end, the community should be decomposed into the most modular units possible. However the designer should ensure that these units are of sufficient granularity to warrant the overhead inherent in goal distribution - distributing small tasks can prove more expensive than performing them in one place (Durfee *et al.*, 1987).

The disadvantage of distributing control and data is that knowledge of the system's overall state is dispersed throughout the community and each individual has only a partial and imprecise perspective. Thus there is an increased degree of uncertainty about each agent's actions, meaning that it more difficult to attain coherent global behaviour - for example, agents may spread misleading and distracting information, multiple agents may compete for unshareable resources simultaneously, agents may unwittingly undo the results of each others activities and the same actions may be carried out redundantly. Also the dynamics of such systems can become extremely complex, giving rise to nonlinear oscillations and chaos (Huberman and Hogg, 1988). In such cases the coordination process becomes correspondingly more difficult as well as more important[1].

To develop better and more integrated models of coordination, and hence improve the efficiency and utility of DAI systems, it is necessary to obtain a deeper understanding of the fundamental concepts which underpin agent interactions. The first step in this analysis is to determine the perspective from which coordination should be described. When viewing agents from a purely behaviouristic (external) perspective, it is, in general, impossible to determine whether they have coordinated their actions. Firstly, actions may be incoherent even if the agents tried to coordinate their behaviour. This may occur, for instance, because their models of each other or of the environment are incorrect. For example, $robot_1$ may see $robot_2$ heading for $exit_2$ and, based on this observation and the subsequent deduction that it will use this exit, decide to use $exit_1$. However if $robot_2$ is heading towards $exit_2$ to pick up a particular item and actually intends to use $exit_1$ then there may be incoherent behaviour (both agents attempting to use the same exit) although there was coordination. Secondly, even if there is coherent action, it may not be as a consequence of coordination. For example imagine a group of people are sitting in a park (Searle, 1990). As a result of a sudden downpour all of them run to a tree in the middle of the park because it is the only available source of shelter. This is uncoordinated behaviour because each person has the intention of stopping themselves from becoming wet and

1. Similar experiences have also been noted in organisational science: the greater the task uncertainty, the greater the amount of information which must be processed among decision makers during task execution in order to achieve a given level of performance (Galbraith, 1973).

even if they are aware of what others are doing and what their goals are, it does not affect their action. This contrasts with the situation in which the people are dancers and the choreography calls for them to converge on a common point (the tree). In this case the individuals are performing exactly the same actions as before, but it is coordinated behaviour because they each have the aim of meeting at the central point as a consequence of the overall aim of executing the dance. For these two reasons, the coordination process is best studied by examining the internal structure of the individual agents (i.e. the agents' beliefs, desires, preferences, intentions, and so on).

Having decided upon a perspective, the next decisions concern the model that will be used to describe the problem and the structures that will be used to describe the agents. Here a distributed goal search formalism is used to characterise DAI systems (section 2) and the key agent structures are *commitment* and *convention* (section three). This model of coordination is founded upon the "*Centrality of Commitments and Conventions Hypothesis*" which states that: *all coordination mechanisms can ultimately be reduced to commitments and their associated (social) conventions*. Commitments are viewed as pledges to undertake a specified course of action, while conventions provide a means of monitoring commitments in changing circumstances. The former provide a degree of predictability so that agents can take the (future) activities of others into consideration when dealing with inter-agent dependencies, global constraints or resource utilization conflicts. The latter provide the flexibility which cooperating agents need if they are to cope with being situated in dynamic environments. To operate effectively when the external world and their own beliefs are constantly changing, agents must possess a mechanism for evaluating whether existing commitments are still valid. Conventions provide this mechanism: defining the conditions under which commitments should be reassessed and specifying the associated actions which should be undertaken in such situations. Finally, section four investigates three prominent coordination techniques (organisational structuring, meta-level information exchange and multi-agent planning) and shows how they can all be reformulated in terms of commitments and conventions - thus providing further evidence for the main claim of this paper.

2. Modelling Distributed AI Systems as a Distributed Goal Search Problem

Several authors have recently characterised DAI as a form of distributed goal search with multiple loci of control (Durfee and Montgomery, 1991; Gasser, 1992; Jennings, 1993; Lesser, 1991). Adopting Lesser's basic formalism, the actions of $Agent_1$ and $Agent_2$ in solving goals $G^1{}_0$ and $G^2{}_0$ respectively can be expressed as a classical AND/OR goal structure search[2] (figure 1). The classical graph structure has been augmented to include a representation of the interdependencies between the goals and to indicate the resources needed to solve the primitive goals (leaf nodes). Interdependencies can exist between high level sibling goals, such as $G^1{}_1$ and $G^1{}_2$, or they can be more distant in the goal structure (eg between $G^1{}_{1,1}$ and $G^2{}_{p,2}$). In the latter case, $G^1{}_1$ and $G^2{}_p$ become interacting goals if $G^1{}_{1,1}$ is used to solve $G^1{}_1$. Indirect dependencies exist between goals through shared resources (eg $G^1{}_{m,1,2}$ and $G^2{}_{p,2,2}$ through resource $d^1{}_j$). Resource dependencies can be removed simply by providing more of the resource in question; dependencies between goals, on the other hand, cannot be circumvented as they are a logical consequence of the community's environment. In all other aspects, the two types of dependency are identical.

2. Figure 1 represents a typical multi-agent situation in which each individual has its own goals, but it must interact with others to achieve them. In contrast, a distributed problem solving system would have a single root node corresponding to the common objective.

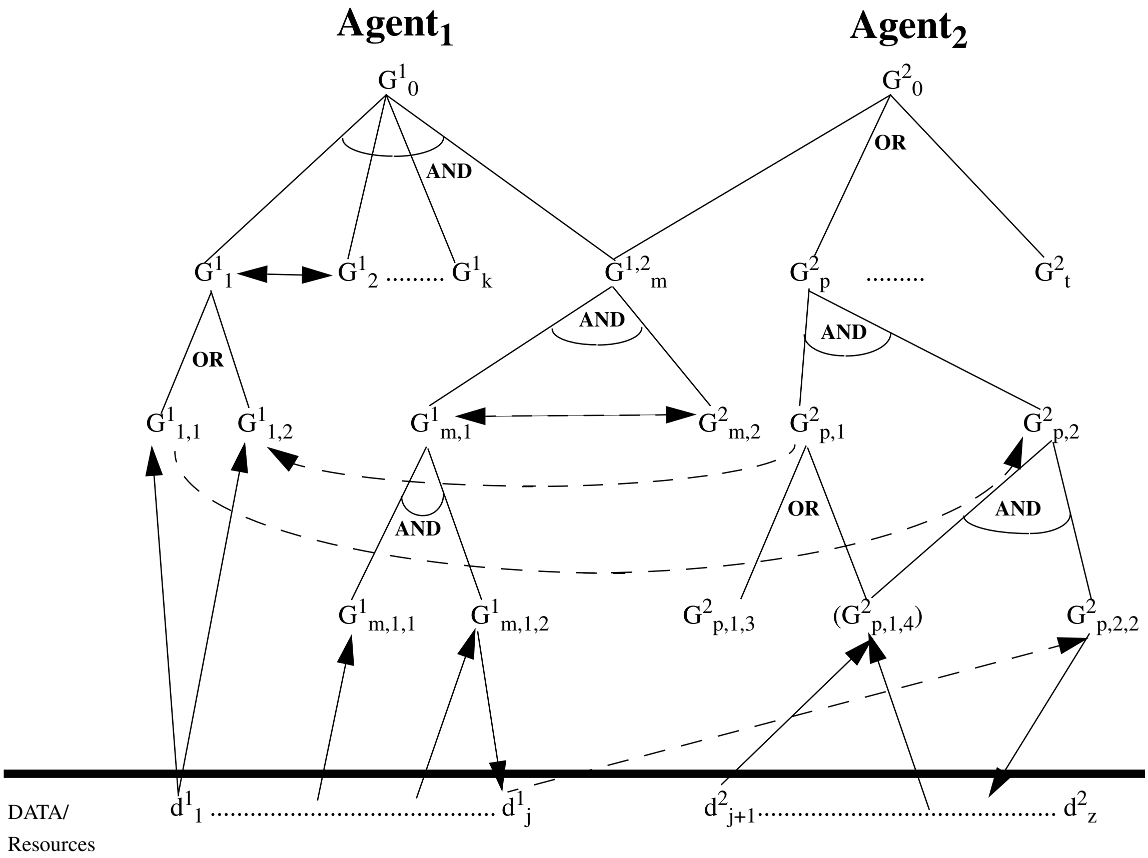

Figure 1: A distributed goal search tree involving $Agent_1$ and $Agent_2$. The dotted arrows indicate interdependencies between goals and data in different agents, solid arrows dependencies within an agent. The superscripts associated with goals and data indicate the agent which contains them (Jennings, 1993).

Interdependencies can be classified along two orthogonal dimensions: whether they are weak or strong, and whether they are uni-directional or bi-directional. S*trong dependencies* must be satisfied if the dependent goal is to succeed; *weak dependencies* facilitate or constrain problem solving but need not be fulfilled for the dependent goal to succeed. An example of a strong dependency is where the output of a goal (G) is a mandatory input (I) for the dependent goal (DG) and where G is the only source of I in the community. A weak dependency exists if there is more than one source for I or if I is an optional input for DG. A *uni-directional* dependency (written $G^1_{1,1} \longrightarrow G^2_{p,2}$) means that $agent_2$'s goal $G_{p,2}$ is dependent (either strongly or weakly) on $agent_1$'s goal $G_{1,1}$, but $G^1_{1,1}$ is unaffected by $G^2_{p,2}$; with *bi-directional* dependencies (written $G^1_{m,1} \longleftrightarrow G^2_{m,2}$) the goals of both agents are affected. The provision of information I by goal G for DG is an example of a uni-directional dependency ($G \longrightarrow DG$); a bi-directional dependence occurs, for example, when two goals need to be performed simultaneously.

The nature of the inter-agent dependencies is the critical determinant of the type of coordination which will take place. For example, if $Agent_1$ knows that $G^2_{p,2,2}$ requires resource d^1_j before it can start (strong dependency, uni-directional), then it may decide to execute $G^1_{m,1,2}$ (to produce the necessary resource) before $G^1_{m,1,1}$ if there is no other information distinguishing between these two alternatives. Secondly, the relationship between $G^1_{m,1}$ and $G^2_{m,2}$ may stipulate that both actions need to be performed simultaneously (strong dependency, bi-directional) in which case the two agents need to reach an agreement about the respective execution times. Finally, if $Agent_1$ chose $G^1_{1,1}$ as a means of satisfying G^1_1 the result of this task may provide valuable information (weak dependency, uni-directional) which $Agent_2$ could use when solving $G^2_{p,2}$ (eg it may provide a partial result which enables $G^2_{p,2}$ to be significantly shorter). Knowing this, $Agent_1$ will invoke an information sharing form of cooperation to supply $Agent_2$ with the necessary result when it becomes available.

It was necessary to extend Lesser's graph formalism to allow joint goals (eg $G^{1,2}{}_{m}$) because joint goals are the basis of joint action (i.e. there can be no joint action unless there is first a joint goal). Joint actions are a sophisticated form of cooperation in which a team of agents decide to pursue a common goal in a cooperative manner (this contrasts with simpler forms of cooperation such as asking an agent to perform a single task or spontaneously volunteering relevant information to interested acquaintances). This form of interaction can be characterised as having the following properties: (i) the team members are mutually responsive to one another, (ii) the team members have a joint commitment to the joint activity, and (iii) the team members are committed to be mutually supportive of one another during the pursuit of their joint objective (Bratman, 1992). Joint goals differ from individual goals in that they are not directly associated with actions - for this reason, they must be mapped onto individual goals as only individual agents have the ability to act (perform domain level tasks). However joint goals can be in the mind of each individual which is acting as part of the collective, implying that everything necessary for team behaviour can be possessed by individual agents, even though the aim makes reference to the collective. Thus the joint goal $G^{1,2}{}_{m}$ is internalised within $Agent_1$ and $Agent_2$ and results in $Agent_1$ performing $G^{1}{}_{m,1}$ and $Agent_2$ performing $G^{2}{}_{m,2}$.

Lesser (1991) makes the following general observations about the graph formalism. The entire goal structure need not be fully elaborated in order for problem solving to begin, it may be constructed as problem solving progresses. Developing the graph can be a complex social activity involving negotiation, persuasion and the resolution of conflicts or it may be undertaken centrally by one agent. Construction can involve a top-down elaboration based on higher-level goals, a bottom-up process driven by the data, or a mixture of the two. Finally, the formalism says nothing about whether the structure is statically defined or evolves dynamically from a composite view of the current, local goal structures of the individual agents.

Formulating a multi-agent system in this manner allows the activities which may require coordination to be clearly identified. Such activities include: (i) defining the goal graph (including identification and classification of interdependencies); (ii) assigning particular regions of the graph to appropriate agents; (iii) controlling decisions about which areas of the graph to explore; (iv) traversing the goal structure; and (v) ensuring that successful traversal of the search space is reported. Some of these activities may be collaborative and some may be carried out by an individual acting in isolation. Determining the approach adopted for each of the various phases is a matter of system design. It will depend upon the nature of the domain (eg in applications in which agents have distinct expertise, assignment of goals simply becomes a matter of identifying the individual capable of performing the activity), the type of agents included in the community (eg with autonomous agents, the global search space is given by the union of the local search spaces and each agent works on its own local goals), and the desired solution characteristics (eg to increase the likelihood of an important result being produced, the same area of the search space may be redundantly assigned to multiple agents, whereas if the desire is to optimise agent usage then such an arrangement is inefficient). This paper concentrates on the problems of deciding which areas of the graph to explore, actually executing the goal structure and ensuring that successful traversal of the goal graph is reported.

3. The Commitment and Convention Model of Coordination

This section describes the process of coordination in terms of the distributed goal graph of the previous section and shows that commitments and conventions are the key mechanisms controlling this activity. This section also argues for the centrality of commitments and conventions hypothesis which states that:

> ***All coordination mechanisms can ultimately be reduced to commitments and their associated conventions***
>
> (Jennings, 1993)

3.1 Detailing Commitments and Conventions

This subsection provides a more precise characterisation of the properties of commitments (section 3.1.1), joint commitments (section 3.1.2), conventions (section 3.1.3) and social conventions (section 3.1.4) before section 3.2 shows how they are the key to the coordination process.

3.1.1 Commitments

Commitments can be seen to have a number of important properties (Becker, 1960; Bond, 1989; Bratman, 1987; Dennett, 1987; Fikes, 1982; Searle, 1983). Agents can make pledges about both actions and beliefs and these pledges can either be about the future or the past. Thus agent A can commit itself to play cricket tomorrow (object of commitment = action, time = future) and agent B can commit itself to believe a particular version of events about the reasons for the start of World War I (object of commitment = belief, time = past). For the purposes of coordination, however, the most important commitments are related to present and future beliefs and actions. No fundamental differences between pledges which are internalised within an agent (eg I will lose 12 pounds in weight) and pledges which are made to a second party (eg I will fix your car for you) are assumed. Commitments may be conditional - for example, A will play cricket tomorrow if the weather is sunny. Finally, a pledge to undertake an activity involves an associated commitment about the resources required to carry out that action. If A pledges to play cricket tomorrow, then it is also devoting its resources of time and energy to this activity.

If an agent commits itself to perform a particular action then, provided that its circumstances do not change, it will endeavour to honour that pledge. This obligation constrains an agent’s subsequent decisions about undertaking fresh activities since it knows that sufficient resources must be reserved to honour its existing commitments. If an agent had infinite resources which could be freely allocated to any permutation of its commitments then there would be no such restrictions. However as most resources are finite, and constraints are often imposed by the environment, an agent is limited in the number and type of commitments it can make. For this reason, an agent’s commitments should, as far as it is aware, be both internally consistent and consistent with its beliefs.

Commitments can be made at many different levels and have correspondingly different time horizons. When $Agent_1$ pledges to perform $G^1{}_0$ this will invariably be a high level objective (eg diagnose faults in an electricity network) to which it will probably remain committed for some considerable amount of time. The leaf nodes, on the other hand, will involve fairly specific courses of action (eg see if there is a fault in low voltage $line_1$) and will have a correspondingly shorter duration.

Generally, the greater the degree of accuracy to which an agent knows its acquaintances' commitments, the more detailed its predictions can be and so the more coherently the community will behave. However it is not always desirable to transmit all of the low-level details about commitments - rather agents should communicate at a sufficiently detailed level to promote satisfactory coordination, but at a sufficiently abstract level to ensure agents retain a degree of flexibility in achieving their objectives in an uncertain environment. For example, knowing that $Agent_2$ is committed to G^2_0, gives no indication of whether $G^2_{p,1}$ will be performed. However knowing that $Agent_2$ is committed to G^2_p means that it is possible to predict that $G^2_{p,1}$ will indeed be performed and that $Agent_1$ can delay its processing of $G^1_{1,2}$ to benefit from the weak, uni-directional dependence. Even if $Agent_2$ communicated a more detailed description of its commitments, for instance that it will perform $G^2_{p,1,3}$, this information will be of no additional benefit to $Agent_1$ since it is not dependent on how $G^2_{p,1}$ is achieved. Not sending details of how $G^2_{p,1}$ will be achieved also leaves $Agent_2$ unconstrained as to whether it will use $G^2_{p,1,3}$ or $G^2_{p,1,4}$.

3.1.2 Joint Commitments

When agents decide to pursue a joint action, they must jointly commit themselves to a joint goal which will bring about the desired state of affairs (Cohen and Levesque, 1991; Grosz and Sidner, 1990; Jennings, 1992; Kinny *et al.*, 1992; Rao *et al.*, 1992; Searle, 1990; Tuomela and Miller, 1988). This joint commitment has all the aforementioned properties of individual commitment, but it has the additional constraint that it involves more than one agent[3]. This means the overall state of the joint commitment is distributed. (In contrast, with individual commitments the agent which has made the pledge is aware of its exact status as it forms part of its internal state). So, for example, the state of the joint commitment to the joint goal $G^{1,2}_m$ is distributed between $Agent_1$ in its processing of $G^1_{m,1}$ and $Agent_2$ in its processing of $G^2_{m,2}$. Ideally all team members should have access to a shared mental state related to the joint commitment as this would ensure that they all have the same experiences and beliefs simultaneously. In such cases there would be no divergence amongst the group's members. However since joint actions are undertaken by individuals, and not the team *en masse*, it is the individuals which have first exposure to events related to the joint commitment. Thus a shared mental state is impossible, unless all the agents possess a single common structure which records all of their beliefs about the joint commitment (i.e. agents cannot have any local or private beliefs about the joint action). For example in a team search, if one agent satisfies the group's objective by finding the target item, then at that precise instant in time it is the only one which knows that the joint commitment has been fulfilled. This agent may subsequently inform the others of its achievements, meaning they all share a common perspective once more; however, in the meantime, the cooperating agents have diverged in their beliefs about the joint commitment.

3.1.3 Conventions

An agent should honour its commitments provided that its circumstances do not change. However in the majority of realistic scenarios agents are situated in time-varying and unpredictable contexts - the external world may change, the agent may become aware of new information, another agent may attempt to interact with it, and so on (Jennings, 1994). Therefore in many cases an agent's beliefs and desires will change between the making of a commitment and the associated intention actually being performed - in fact, the longer the time between these two events, the greater the likelihood of a change occurring. In some instances

[3]. A joint commitment involving one agent is equivalent to an individual commitment.

these changes will leave the agent's commitments unaffected, but in other cases commitments may need to be reviewed. For example, if agent A is informed that the first customer at a new garage opening tomorrow will receive a Ferrari, then it may indeed revise its commitments about playing cricket. Therefore commitments should be relatively stable over time, but they should not be irrevocable.

To operate successfully and intelligently, agents need general policies for governing the reconsideration of their commitments. These *conventions* describe circumstances under which an agent should reconsider its commitments. They also indicate the appropriate course of action to either retain, rectify or abandon these commitments. An agent may have several different conventions at its disposal, although each of its commitments should be tracked using precisely one convention.

When specifying conventions a balance needs to be reached between constantly reconsidering all commitments (which will enable the agent to respond rapidly to changing circumstances, but will mean that it spends a significant percentage of its time reasoning about action rather than actually carrying out useful tasks) and never reconsidering commitments (which means agents spend most of their time acting, but what they are actually doing may not be particularly relevant in the light of subsequent changes). Kinny and Georgeff (1991) carried out a series of experiments in which different conventions were examined in environments exhibiting different rates of change. In all cases it was found that the "bold" agents (those which never reconsidered their commitments) performed better than the "normal" agents (those which are slightly more open to reconsideration) which were better than the "cautious" agents (those which were prone to reconsideration). However in rapidly changing and uncertain contexts, the utility of a relatively sophisticated convention is significantly increased. Indeed empirical evaluation has shown that in such circumstances conventions play a pivotal role in ensuring the community acts in a coherent manner (Jennings and Mamdani, 1992).

The lists of both the situations under which commitments should be reassessed and the actions which should be taken in such circumstances can be empty. So an agent can remain permanently committed to a goal even if it has been achieved or an agent can take no action as a result of changes in its circumstances. By means of an illustration, figure 2 shows a convention based upon Cohen and Levesque's (1990) model of individual rational behaviour - it shows that an agent can renege upon a commitment if it believes that the commitment is satisfied or unattainable or if the motivation for the commitment is no longer present.

```
CONVENTION: Cohen and Levesque Model

REASONS FOR RE-ASSESSING COMMITMENT:
          • COMMITMENT SATISFIED
          • COMMITMENT UNATTAINABLE
          • MOTIVATION FOR COMMITMENT NO LONGER PRESENT

ACTIONS:

   R1:  IF   COMMITMENT SATISFIED OR
             COMMITMENT UNATTAINABLE OR
             MOTIVATION FOR COMMITMENT NO LONGER PRESENT
        THEN DROP COMMITMENT
```

Figure 2: Sample Convention

3.1.4 Social Conventions

Although conventions play an important role in DAI systems, they are essentially asocial constructs. They describe how an agent should monitor its commitments, but they do not specify how an agent should behave towards its fellow community members if it alters or modifies its commitments. For goals which are unrelated to other activities this stance is sufficient. However for goals which are inter-dependent, it is essential that the relevant acquaintances are informed of any substantial change which affects their processing, if the community is to act in a coherent manner. For this reason, agents need social conventions which specify how to behave with respect to the other community members when their commitments alter. Designing such conventions is a skilful activity. On the one hand it is important that relevant information pertaining to changes in commitment is disseminated at the earliest possible opportunity; on the other hand, agents should not broadcast information about their commitments each and every time they change because this will over burden the communication resources and needlessly distract the recipients. A sample social convention which reflects this trade-off is shown in figure 3.

```
SOCIAL CONVENTION: Limited-Bandwidth

INVOKE WHEN:
            • LOCAL COMMITMENT DROPPED
            • LOCAL COMMITMENT SATISFIED
ACTIONS:

   R1:  IF LOCAL COMMITMENT SATISFIED
        THEN INFORM ALL RELATED COMMITMENTS

   R2:  IF LOCAL COMMITMENT DROPPED BECAUSE UNATTAINABLE OR MOTIVATION
                 NOT PRESENT
        THEN INFORM ALL STRONGLY RELATED COMMITMENTS

   R3:  IF LOCAL COMMITMENT DROPPED BECAUSE UNATTAINABLE OR MOTIVATION
                 NOT PRESENT
           AND COMMUNICATION RESOURCES NOT OVERBURDENED
        THEN INFORM ALL WEAKLY RELATED COMMITMENTS
```

Figure 3: Sample Social Convention

Ideally the participants should be mutually aware of the convention which governs their interaction. Such awareness is needed if the agents are to minimise the uncertainty in their collaboration and maximise the benefit of the joint action. Thus, for example, if $Agent_2$ must have resource $d^1{}_j$ to perform $G^2{}_{p,2,2}$ then it will invoke a task sharing form of cooperation and ask $Agent_1$ to make this resource available. However merely asking for $d^1{}_j$ to be produced is not sufficient because $Agent_2$ also wants to be informed when it is available. To ensure the necessary dissemination occurs, $Agent_2$ must request that the resource be produced using an appropriate social convention (eg that of figure 3). Whether $Agent_1$ accepts this proposal will depend upon its personal preferences and the relative authority of the two agents. If the proposal is acceptable, or if $Agent_2$ can force $Agent_1$ to use it, then the convention will be adopted. If the proposal is unacceptable, then the two agents will have to enter a negotiation phase to decide upon an acceptable solution. Alternatively, rather than having to determine the social convention for each and every interdependent goal at run time, which will significantly slow down processing, the system designer may stipulate that when two agents interact they will always use a particular convention. He may even specify that the whole community must use a particular convention for all their joint actions.

The social conventions specified up to this point have all been concerned with inter-related individual goals. However all joint actions must have an overarching joint goal and a corresponding joint commitment. This begs the question: how does the social convention for a joint commitment differ from the social convention of an individual commitment which is related to other goals? The answer to this question lies in the mutual support requirement of joint actions. The minimum form of mutual support that a team of cooperating agents should provide to one another is to share information about: (i) the status of their commitment to the shared objective; (ii) the status of their commitment to the given team framework[4]. If an agent's beliefs about either of these issues change, then it is part of the intuitive semantics of joint commitments that all team members are informed. As many joint actions depend upon the participation of all their team members, a change of commitment by one participant can also jeopardise the whole group's efforts. Hence if an agent comes to believe that a fellow team member is no longer jointly committed, it also needs to reassess its own position with respect to the joint action. These three basic assumptions are encoded in a convention which represents the minimum state of affairs for joint commitments (figure 4). Thus whereas any convention can be used for individual commitments (even for those which have related goals), joint commitments require each team member to adhere to the basic joint action convention. In some applications it may be desirable to have more sophisticated joint action conventions, but in every case they must still incorporate these fundamental ideals.

```
BASIC JOINT ACTION CONVENTION

INVOKE WHEN:
          • STATUS OF COMMITMENT TO JOINT GOAL CHANGES
          • STATUS OF COMMITMENT TO ATTAINING JOINT ACTION IN
                 PRESENT TEAM CONTEXT CHANGES
          • STATUS OF JOINT COMMITMENT OF A TEAM MEMBER CHANGES

ACTIONS:

   R1:   IF STATUS OF COMMITMENT TO JOINT ACTION CHANGES OR
            STATUS OF COMMITMENT TO PRESENT TEAM CONTEXT CHANGES
         THEN INFORM ALL OTHER TEAM MEMBERS OF CHANGE

   R2:   IF STATUS OF JOINT COMMITMENT OF A TEAM MEMBER CHANGES
         THEN DETERMINE WHETHER JOINT COMMITMENT STILL VIABLE
```

Figure 4: Basic Joint Action Convention

3.2 Commitments and Conventions: The Cornerstones of Coordination

Durfee *et al.*, (1989) identify three major ingredients which must be present for successful coordination: (i) there must be structures which enable the agents to interact in predictable

[4] This stipulation covers the situation in which an agent which is initially committed to the joint action decides to leave the team, but continues to pursue the joint objective in an individualistic manner. Detecting this situation is important because the agent which is no longer committed to the team framework will follow its own solution path, without considering its effects on those remaining in the original team.

ways; (ii) there must be flexibility so that agents can operate in dynamic environments and can cope with their inherently partial and imprecise viewpoint of the community; and (iii) the agents must have sufficient knowledge and reasoning capabilities to exploit the available structure and the flexibility. The final component, which they fail to explicitly mention, is the necessity of structures to provide mutual support to the cooperating agents. In the remainder of this subsection it is shown that: commitments provide the necessary structure for predictable interactions, conventions provide the flexibility needed to operate in dynamic environments, and social conventions provide the necessary degree of mutual support. Thus,

***Coordination* = *Commitments* +**
***Conventions* +**
***Social Conventions* +**
Local Reasoning

These assertions are demonstrated by examining the multifarious social interactions which occur as a consequence of the various types of goal relationships and goal interdependencies. For each distinct case the central role of commitments, conventions and social conventions is highlighted.

3.2.1 Commitments in Goal-Subgoal Relationships

If the goal being undertaken by a group of cooperating agents involves an “AND” relationship, then all commitments to the sub-goals must be honoured if the parent goal is to succeed. If just one agent reneges, then the other agents’ activities are doomed in their present form. Therefore it is only the belief that others will honour their commitments that makes it rational for an agent to carry out its part. Without this confidence no agent would carry out its individual processing, since achievement of the sub-goals in isolation is unlikely to bring it any benefits. Hence AND goals cannot be achieved by cooperative problem solving without the notion of commitment.

If a goal is composed of a number of sub-goals, organised in an “OR” relationship, and these are each assigned to different agents which all carry out their activities in parallel, then failure of one agent to fulfill its commitment will not jeopardise achievement of the parent goal. However if all the agents renege upon their obligations then the parent goal will not be achieved; thus the commitment of at least one agent must be guaranteed. If agents coordinate their activities more closely, and arrange for only one of the alternate sub-goals to be carried out at any one time (to avoid needless duplication), then commitment failure can have serious repercussions for the community’s overall level of coherence. For example, if a team agrees that $Agent_1$ will carry out the subgoal which fulfills the parent goal G, then the remaining agents can continue with their processing and can make subsequent commitments based on the fact that G will indeed be achieved and that they do not have to expend resources towards this end. However if $Agent_1$ does not honour its pledge, then provided that G is still desired, one of the other agents will have to carry out some unexpected processing activity. This additional work may conflict with commitments which the agent has subsequently made and may result in it having to delay, or even abandon, some of them because additional resources are unexpectedly required to achieve G. Such delays may have a knock-on affect to other agents, causing the community to operate ineffectively and requiring it to undertake a significant amount of replanning.

3.2.2 Commitments in Goal Dependency Relationships

Consider the situation in which $Agent_1$ and $Agent_2$ have the respective interrelated goals $G^1{}_1$ and $G^2{}_1$. If there is a strong bi-directional dependence then both agents must honour their commitments, otherwise neither of them will be able to achieve their objectives. If the relation is strong but uni-directional ($G^1{}_1 \longrightarrow G^2{}_1$), then failure of $Agent_1$ to honour its commitment means that $Agent_2$ will be unable to achieve $G^2{}_1$ and it will either have to find an alternative path for achieving the parent goal or abandon it completely.

With weak dependencies, the agents involved may still be able to proceed but this may be in a suboptimal manner. For example, an agent may have delayed processing an action on the premise that an acquaintance will provide it with sufficient information to significantly speed up its problem solving. If this information is no longer forthcoming, because the acquaintance changed its commitments, then the agent has wasted potentially useful processing time. As another example, an agent may select a certain path through the search space in the belief that information which will be provided through a weak dependency will make this path less expensive than its alternatives. But if the agent providing the information reneges upon its commitment and the information is not forthcoming, then the chosen path may be suboptimal.

If the relationship is bi-directional and one agent fails to fulfill its pledge, the agent which is still committed to its side of the bargain may be adversely affected since it chose to undertake the goal believing that it would be able to profit from the commitment of the other agent. In the uni-directional case ($G^1{}_1 \longrightarrow G^2{}_1$) if $Agent_1$ changes its mind, $Agent_2$'s processing of $G^2{}_1$ will be adversely affected for the reason described above.

In both the weak and the strong uni-directional cases, if $Agent_2$ drops its commitment to $G^2{}_1$ then this may have a detrimental effect on $Agent_1$. This is the case if $Agent_1$ chose $G^1{}_1$, even though it was locally suboptimal, because the net utility to the community of the performance of the pair $\{G^1{}_1, G^2{}_1\}$ whilst satisfying the specified relationship was higher than if $Agent_1$ chose an alternative to $G^1{}_1$ and $Agent_2$ chose $G^2{}_1$. However the potential benefits of $Agent_1$'s sacrifice were not observed because $Agent_2$ failed to carry out its pledge about $G^2{}_1$.

3.2.3 Social Conventions in Goal-Subgoal Relationships

Social conventions report changes in commitments to dependent agents and are especially important when there is an “AND” relationship between a goal and its constituent sub-goals. Without adequate information dissemination the other agents will remain committed to performing their sub-goals even though they will not satisfy their original purpose of fulfilling the parent goal. This is also the case for “OR” relationships in which only one of the sub-goals is active at any one time. Adherence to a suitable convention provides a secondary degree of confidence in joint actions - the group believes that each individual will do their best to perform their subpart, but if they are unable to keep their commitment then they must inform their fellow team members of their change in state.

3.2.4 Social Conventions in Goal Dependency Relationships

In terms of interagent goal dependencies ($G^1{}_1 \longrightarrow G^2{}_1$), reports of changes in the status of commitments are essential if the relationship is strong and bi-directional, and if $Agent_1$ reneges on its goal and the relation is strong and uni-directional. In all other cases reports on changes in commitments are desirable in that they may enable the agent which is still committed to reassess its position. This may result in it choosing a different path through the graph - either because it is freed from the constraint of having to honour the relationship or because it can no longer benefit from the interaction with the agent which is no longer committed.

3.2.5 The Basic Joint Action Convention, Joint Commitments and Joint Goals

With respect to joint actions, the most important feature of joint commitments and social conventions is that they enable individuals to make assumptions about the actions of other community members. They provide a degree of predictability to counteract the uncertainty caused by the distribution of control. So for the joint goal $G^{1,2}{}_{m}$, $Agent_2$ can carry out $G^{2}{}_{m,2}$ in the knowledge that $Agent_1$ is probably performing $G^{1}{}_{m,1}$ and that, if it is not, then it will at least be trying to inform it of this change (because of the basic joint action convention). Without this assurance there would be no point in $Agent_2$ even starting $G^{2}{}_{m,2}$ since it is only carrying out this activity to achieve the joint goal and the joint goal requires both sub-goals to be fulfilled. Thus each agent is only carrying out its respective action because it believes that the other is also doing its bit.

4. Common Coordination Techniques and their use of Commitments and Conventions

In this section three of the most common mechanisms for managing the coordination process in DAI systems are presented and analysed - these include organisational structuring (section 4.1), exchanging meta-level information (section 4.2), and multi-agent planning (section 4.3). For each technique there is a brief statement about how it facilitates the coordination of behaviour, what its main characteristics are and how it can be reformulated in terms of the centrality of commitments and conventions hypothesis.

4.1 Organisational Structures

In the context of DAI systems, an organisational structure can be viewed as a pattern of information and control relationships between individuals (Gasser, 1992). These control relationships, be they hierarchical, heterarchical or flat, are responsible for designating the relative authority of the agents and for shaping the types of social interaction which can occur. For example, when building a community of agents for diagnosing faults in an electricity network (Cockburn and Jennings, 1995), the system designer may specify a functional organisation ($agent_1$ works on high voltage faults, while $agent_2$ works at the low voltage level) or a spatial organisation ($agent_1$ deals with all types of faults in $region_1$, $agent_2$ with all types of faults in $region_2$). Concentrating on the spatial distribution, and assuming the interest areas of the agents overlap, the authority relationships determine how redundances are avoided - in a hierarchy, high level nodes inform the lower level ones of the activities they are to pursue; whereas in a flat structure this process is only achievable through direct negotiation between the parties concerned.

The relationships specified by the organisational structures give general, long-term information about the agents and the community as a whole. They aid the coordination process by specifying which actions an individual will undertake and by providing a means of dividing up the search space without having to go into detail about the particular sub-trees. Other authors have followed this basic approach using different terminology - Singh (1990) employs the notion of “strategies” to provide an abstract specification of the behaviour of an agent or a group and Werner (1989) uses “roles” for describing expectations about individual behaviour. Shoham and Tennenholtz (1992) propose a more detailed organisational form which they term a “social law”. With this approach, the society adopts a set of laws (eg road traffic rules) which specify how individuals should behave. Each programmer is then committed to obeying these laws when building his individual agent. The design process is simplified because it can be assumed that all the other agents will adhere to the specified law.

As a specific illustration of this approach consider the Distributed Vehicle Monitoring Testbed (Decker, 1995) which simulates a spatially organised community of agents performing a distributed interpretation to track vehicles moving amongst them. Each agent decides which areas of the search space to explore based upon its current local view and its organisational knowledge of its role in the community. In this context, coordination consists of two concurrent activities: the construction and maintenance of a community-wide organisational structure, and the continuous elaboration of this structure into precise activities using the local knowledge and control capabilities of each agent. The organisation itself is specified as a set of "interest areas" and as a set of priority ratings. The former indicate what, when and to whom information should be sent; the latter indicate how to evaluate the importance of processing different types of goals. The authority relationships indicate the relative priorities which should be attached to processing externally generated goals versus local goals.

When an agent undertakes a particular role within an organisation it is, in fact, making a high-level commitment about the types of activity it will pursue. For instance in the electricity management scenario, if $agent_1$ undertakes the role of diagnosing high-voltage faults, other agents will expect it to undertake work in this area. They will make subsequent decisions in their local problem solving based on the assumption that $agent_1$ will indeed be dealing with all the faults on the high voltage network.

Although they are relatively long-term structures, it has been shown that different organisations are appropriate for different problem situations and performance requirements (Malone, 1987). Hence as a situation evolves, the community may need to periodically reassess its structure to determine whether it is still appropriate or whether a rearrangement would be beneficial - see the work of Ishida *et al.* (1990) for an illustration of the dynamic reorganisation of a group of cooperating agents in response to changes in the environment. In the electricity management scenario, for example, the community may decide that it is best to replace the agent carrying out high-voltage diagnosis with several spatially distributed agents so that the load and the reliance upon any one individual is reduced. This evaluation corresponds to a convention for the organisational structure.

4.2 Meta-Level Information Exchange

Meta-level information exchange involves agents sending each other control level information about their current priorities and focus (Gasser, 1992). For example in the functionally distributed electricity management scenario described in the previous subsection, $agent_1$, which is working on the low voltage network, may indicate that it believes the most important fault is in $region_1$. Upon receiving this information, $agent_2$, which is working on the high-voltage network, may also decide to concentrate its efforts on this region to determine whether the fault being experienced on the low voltage network is in fact a manifestation of a problem with the high voltage system (eg no supply is getting through).

Durfee (1995) has developed a meta-level information exchange mechanism, called Partial Global Planning, in which agents build and share local plans as a means of identifying potential improvements to coordination. These partial global plans (PGPs) are exchanged by agents as a means of building representations of their acquaintances' activities - they indicate which goals will be pursued and in what order, what results will be achieved and how long each goal is likely to take. Individual community members then use a model of themselves and a representation of their acquaintances to identify when agents have PGPs whose objectives are part of some larger community effort. If such complementary activities are detected, the related PGPs are combined into a single, larger PGP which provides a more complete view of the group's activities. Agents can then revise their other PGPs to reflect the new position. As a specific example, a PGP could indicate that the outcome of one agent's task provides useful predictive information for an acquaintance. This expectation, together with the accompanying transmission of the information, would then be explicitly represented in the PGP, resulting in a plan to use information resources more effectively. As a second example, an agent may survey its current view of community-wide PGPs and identify acquaintances which are being under utilised, whilst there are others which are overburdened. By modifying the relevant PGPs, the agent could propose how the community could transfer subproblems so as to work better as a team.

Meta-level information exchange is a medium term source of knowledge about an agent's commitments - shorter than organisation structures but longer than multi-agent planning approaches (section 4.3). It enhances coordination only to the degree to which it is accurate - indeed inaccurate information may be more detrimental than no information at all. Again it can be seen that once an agent indicates it will work in a particular region of the search space, it is important to honour that commitment - failure to do so will result in misleading information being spread around the network and incoherent problem solving. However, as with the other approaches, commitments should not be irrevocable; some form of convention is needed for monitoring their progression. With the PGP approach, for example, agents often altered their local plans, either because new tasks arrived or because actions took longer than expected, and so their commitments needed to be updated. However if agents informed each other of every minor change in their commitments, it could cause a chain-reaction which spreads throughout the system. Therefore agents adopted an implicit convention in which they informed their acquaintances only when the deviations were deemed significant.

4.3 Multi-Agent Planning

With the multi-agent planning approach to coordination, agents usually form a plan which specifies all their future actions and interactions with respect to achieving a particular objective. It details, before execution commences, the areas of the search space that will be traversed and the route each agent should take at each decision point in the activity[5]. Multi-agent plans are typically built to avoid inconsistent or conflicting actions, particularly with respect to the consumption of scarce resources.

Multi-agent planning differs from organisational structuring and meta-level information exchange in terms of the level of detail to which it specifies every agent's activities. With this approach, agents know in advance exactly what actions they will take, what actions their acquaintances will take and what interactions will occur. By requiring such a complete specification of behaviour, the plans can only realistically have a short time horizon because of problems with the unpredictability and dynamicity of events in the environment. As plan construction has to take into account all the possible choice points the agent would have reached, without the benefit of constraining information from actual execution, this approach often requires substantially more computational and communication resource than the other two mechanisms.

5. Partial global planning differs from multi-agent planning in that the former does not require the agents to reach mutual agreements about all of their activities before they start acting. Indeed because the partial plans can change so fluidly and because it takes time to propagate changes such agreements may never be attained.

There are two basic approaches to multi-agent planning: centralised and distributed. Georgeff (1983) built a system in which the plans of individual agents were developed separately and then sent to a central coordinator which analysed them to identify potential interactions. The coordinator then identified those interactions which had the potential to cause conflicts and grouped the sequences of unsafe actions to create critical regions. Finally, communication primitives were inserted into the individual plans so that the agents synchronised their activities appropriately. Cammarata *et al.* (1983) also devised a centralised multi-agent planning system (for the air traffic control domain). In their system, each aircraft (agent) sends the coordinator information about its intended actions. The coordinator then builds a plan which specifies all of the agents' actions, including the actions that they, or some other node, should take to avoid collisions. With distributed multi-agent planning, the plan is developed by several agents. This means there may be no one individual with a view of the entire community's activities and hence detecting and resolving undesirable interactions becomes significantly more difficult. Corkill (1979) devised a distributed hierarchical planner based on NOAH (Sacerdoti, 1977) where agents represent each other using MODEL nodes and plan execution is coordinated by using explicit synchronisation primitives. Rosenschein and Genesereth (1985) used a logic-based approach to study how agents with a common goal, but different local knowledge, could exchange information to converge on identical plans.

Once a plan has been devised, the agents involved are committed to performing the specified actions. If they believed that their acquaintances are unlikely to keep their pledges, then they would not enter the planning phase in the first place because it is such a resource consuming activity. Commitments are, therefore, the foundation of this approach. There is no latitude for deviation from the agreed course of action because it may introduce resource conflicts or other undesirable side-effects which would impair the community's performance. However the situation may change so radically between generation and execution that if the plan was performed the benefit would be negligible or even negative. In this case it is worth entering a replanning phase to produce a more profitable alternative (Kambhampati and Hendler, 1992; Pollack, 1992). Hence there is a need for conventions to determine when replanning is necessary and whether the existing plan should be reused or whether a fresh plan should be devised.

5. Conclusions

This paper has argued that the process of coordination is built upon four main structures - commitments, conventions, social conventions and local reasoning capabilities. Furthermore it has hypothesised that **all** coordination mechanisms can be expressed using these concepts; to demonstrate this point three of the most common coordination techniques have been reformulated in these terms.

Having posited the key structures it is important that the model is refined in order to more precisely discriminate between a number of different types of social interaction. This refinement should be undertaken by two parallel strands of work: (i) by building DAI applications which explicitly use these concepts (see Huang *et al.*, 1994 for an illustration from the domain of distributed patient care); and (ii) by producing more precise models of some of the key concepts (see Wooldridge and Jennings (1994) for a first formulation of commitments and conventions using a quantified multi-modal logic). This subsequent work should aim to address the following issues:

- Produce a finer grain classification of the types of goal interdependency

• Provide a methodology for designing appropriate conventions and social conventions

• Characterise the process by which social conventions are agreed

• Provide mechanisms which support robust coordination in the face of uncertainty about the social convention which is in operation during a given interaction

• Provide empirical evidence of the effectiveness of different conventions in various environmental and domain circumstances

• Provide a means of characterising the local reasoning which is required to support coordinated behaviour

Acknowledgement

This is a refined and updated version of the paper which appeared in The Knowledge Engineering Review.

References

Becker, H. S., (1960) "Notes on the Concept of Commitment" *American Journal of Sociology* **66** (1) pp 32-40.

Bond, A. H., (1989) "Commitment: Some DAI insights from Symbolic Interactionist Society" *Proc. of 9th Workshop on Distributed AI*, Bellevue, WA, pp 239-261.

Bratman, M. E., (1987) "*Intention, Plans and Practical Reason*" Harvard University Press.

Bratman, M. E., (1992) "Shared Cooperative Activity" *The Philosophical Review* **101** (2) pp 327-341.

Cammarata, S., McArthur, D., and Steeb, R., (1983) "Strategies of Cooperation in Distributed Problem Solving" *Proc. Int. Joint Conf. on AI*, Karlsruhe, Germany, pp 767-770.

Clearwater, S. H., Huberman, B. A., and Hogg, T., (1991) "Cooperative Solution of Constraint Satisfaction Problems" *Science* **254** pp 1181-1184.

Cockburn, D., and Jennings, N. R., (1995) "ARCHON: A Distributed Artificial Intelligence System for Industrial Applications" in this volume.

Cohen, P. R., and Levesque, H. J., (1990) "Intention is Choice with Commitment" *Artificial Intelligence* **42** pp 213-261.

Cohen, P. R., and Levesque, H. J., (1991) "Teamwork" *Noûs* **25** (4) pp 487-512.

Corkill, D. D., (1979) "Hierarchical Planning in a Distributed Environment", *Proc. Sixth Int. Joint Conf. on AI*, Cambridge, MA., pp 168-175.

Corkill, D. D., and Lesser, V. R., (1986) "The Use of Meta-Level Control for Coordination in a Distributed Problem Solving Network" *Proc. Int. Joint Conf. on AI*, Karlsruhe, Germany, pp 748-756.

Decker, K., (1995) "Distributed Artificial Intelligence Testbeds", in this volume.

Dennett, D. C., (1987) "*The Intentional Stance*" MIT Press.

Durfee, E. H., (1995) "Planning in Distributed Artificial Intelligence" in this volume.

Durfee, E. H., Lesser, V. R., and Corkill, D. D., (1987) "Coherent Cooperation among Communicating Problem Solvers" *IEEE Trans. on Computers* **36** pp 1275-1291.

Durfee, E. H., Lesser, V. R., and Corkill, D. D, (1989) "Trends in Cooperative Distributed Problem Solving" *IEEE Trans. on Knowledge and Data Engineering* **1** (1) pp 63-83.

Durfee, E. H., and Montgomery, T. A., (1991) "Coordination as Distributed Search in a Hierarchical Behaviour Space" *IEEE Trans. on Systems Man and Cybernetics* **21** pp 1363-1378.

Fikes, R. E., (1982) "A Commitment-Based Framework for Describing Informal Cooperative Work" *Cognitive Science* **6** pp 331-347.

Galbraith, J., (1973) "*Designing Complex Organizations*" Addison-Wesley.

Gasser, L., (1992) "DAI Approaches to Coordination" in *Distributed Artificial Intelligence: Theory and Praxis* (eds. N. M. Avouris and L. Gasser) Kluwer Academic Publishers pp 31-51.

Georgeff, M. P., (1983) "Communication and Action in Multi-Agent Planning" *Proc. of National Conf. on AI,* Washington, DC, pp 125-129.

Gerson, E. M., (1976) "On Quality of Life" *American Sociological Review* **41** pp 793-806.

Grosz, B. J., and Sidner, C. L., (1990) "Plans for Discourse" in *Intentions in Communication*, (eds P.R.Cohen, J.Morgan and M.E.Pollack), pp 417-444, MIT Press.

Huang, J., Jennings, N. R., and Fox, J., (1994) "Cooperation in Distributed Medical Care" *Proc. Int. Conf. on Cooperative Information Systems*, Toronto, Canada.

Huberman, B. A., and Hogg, T., (1988) "The Behaviour of Computational Ecologies" in *The Ecology of Computation* (ed. B. A. Huberman), pp 77-115, North Holland.

Kambhampati, S., and Hendler, J. A., (1992) "A Validation Structure Based Theory of Plan Modification and Reuse" *Artificial Intelligence* **55** pp 193-258.

Kornfield, W. A., and Hewitt, C. E., (1981) "The Scientific Community Metaphor" *IEEE Trans. on Systems Man and Cybernetics*, **11** (1) pp 24-33.

Ishida, T., Yokoo, M., and Gasser, L., (1990) "An Organisational Approach to Adaptive Production Systems", *Proc of 8th National Conf. on Artificial Intelligence*, Boston, USA, pp 52-58.

Jennings, N. R., (1992) "Towards a Cooperation Knowledge Level for Collaborative Problem Solving" *Proc. 10th European Conf. on AI*, Vienna, Austria, pp 224-228.

Jennings, N. R., (1993) "Commitments and Conventions: The Foundation of Coordination in Multi-Agent Systems" *The Knowledge Engineering Review* **8** (3) pp 223-250.

Jennings, N. R., (1994) "Controlling Cooperative Problem Solving in Industrial Multi-Agent

Systems using Joint Intentions" *Artificial Intelligence*.

Jennings, N. R., and Mamdani, E. H., (1992) "Using Joint Responsibility to Coordinate Collaborative Problem Solving in Dynamic Environments" *Proc. 10th National Conf. on AI*, San Jose, CA., pp 269-275.

Kinny, D. N., and Georgeff, M. P., (1991) "Commitment and Effectiveness of Situated Agents", *Proc. Int. Joint Conf. on AI*, Sydney, Australia, pp 82-88.

Kinny, D., Ljungberg, M., Rao, A., Sonenberg, E., Tidhar, G., and Werner, E., (1992) "Planned Team Activity" *Proceedings of the 4th European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Rome, Italy.

Lesser, V. R., (1991) "A Retrospective View of FA/C Distributed Problem Solving", *IEEE Trans. on Systems Man and Cybernetics* **21** pp 1347-1363.

Lesser, V. R., and Corkill, D. D, (1987) "Distributed Problem Solving" *in Encyclopedia of AI* (ed S. C. Shapiro), pp 245-251, John Wiley and Sons.

Malone, T. W., (1987) "Modelling Coordination in Organizations and Markets" *Management Science* **33** pp 1317-1332.

Pollack, M. E., (1992) "The Uses of Plans" *Artificial Intelligence* **57** pp 43-68.

Rao, A. S., Georgeff, M. P., and Sonenberg, E. A., (1992) "Social Plans: A Preliminary Report", in *Decentralised A I 3* (eds. E. Werner and Y. Demazeau), North Holland, pp 57-76.

Rosenschein, J. S., and Genesereth, M. R., (1985) "Deals among Rational Agents", *Proc. 9th Int. Joint Conf. on Artificial Intelligence*, Los Angeles, USA, pp 91-99

Sacerdoti, E. D., (1977) "*A Structure for Plans and Behaviour*" Elsevier.

Searle, J. R., (1983) "*Intentionality: An Essay in the Philosophy of Mind*" Cambridge University Press.

Searle, J. R., (1990) "Collective Intentions and Actions" in *Intentions in Communication*, (Eds P. R. Cohen, J. Morgan and M. E. Pollack), pp 401-416, MIT Press

Shoham, Y., and Tennenholtz, M., (1992) "On the Synthesis of Useful Social Laws for Artificial Agent Societies", *Proc of 10th National Conf. on AI*, San Jose, USA, pp 276-28.

Singh, M. P., (1990) "Group Intentions" *Proc. of 10th Int. Workshop on Distributed AI,* Texas, MCC Technical Report ACT-AI-355-90.

Tuomela, R., and Miller, K., (1988) "We-Intentions" *Philosophical Studies* **53** pp 367-389.

Werner, E., (1989) "Cooperating Agents: A Unified Theory of Communication and Social Structure" in *Distributed Artificial Intelligence Vol II* (eds. L. Gasser and M. N. Huhns), pp 3-36.

Wooldridge, M. J., and Jennings, N. R., (1994) "The Cooperative Problem Solving Process: A Formal Model" Knowledge Engineering Applications Group Technical Report, Dept. Electronic Engineering, Queen Mary & Westfield College.