USING JOINT INTENTIONS IN ELECTRICITY TRANSPORT MANAGEMENT

Nick Jennings
Dept. Electronic Engineering, Queen Mary & Westfield College
Mile End Rd., London E1 4NS
nickj@qmw.ac.uk

1. Introduction

There have been few serious attempts at applying Distributed AI (DAI) techniques to real size, industrial problems. One of the major stumbling blocks to this advancement has been the lack of a clear, implementable theory describing how groups of agents should interact during collaborative problem solving. Such a theory becomes especially important in complex domains in which events occur at unpredictable times, in which decisions are based on incomplete and imprecise information, in which agents possess multiple areas of problem solving competence and when social interactions are complex (i.e. involve several iterations over a prolonged period of time). In such environments it is difficult to ensure that a group's behaviour remains coordinated, because initial assumptions and deductions may be incorrect or inappropriate; therefore a comprehensive theory must provide a grounded basis from which robust problem solving communities can be constructed.

Intentions, a commitment to present and future plans, are considered to be essential in guiding the actions of an individual. However to describe the actions of a *group of agents* working collaboratively the notion of joint intentions, a joint commitment to perform a collective action while in a certain shared mental state, is needed to bind the actions of team members together. Most accounts concentrate exclusively on what it means for a joint intention to exist; this description being in terms of nested structures of belief and mutual belief about the goals and intentions of other agents within the community. In contrast, the notion of *joint responsibility*[1] outlined here stresses the role of intentions as conduct controllers - specifying how agents should behave whilst engaged in collaborative problem solving. This behavioural specification offers a clearer path from theory to implementation; provides functional guidelines for architecture design, criteria against which the monitoring component can evaluate ongoing problem solving and a prescription of how to act when collaborative problem solving becomes untenable. The responsibility model has been implemented and demonstrated on the exemplar domain of monitoring electricity transport networks[2]. The problems faced in this domain are typical of many industrial applications - especially the need to respond to the dynamics of the process being controlled/monitored and taking decisions using partial, imprecise views of the system.

2. Monitoring Electricity Transport Networks

Electricity transportation is concerned with the process of taking electrical energy from where it is produced to where it is consumed. It requires sophisticated monitoring and any problems need to be identified at the earliest opportunity. The CSI (Control System Interface) receives messages from the network and analyses them to determine whether they represent a fault. The AAA (Alarm Analysis Agent) pinpoints elements at fault and the BAI (Blackout Area Identifier) indicates groups of elements out of service (BOA). In the cooperative scenario depicted by fig. 1, the CSI receives an indication that a fault has occurred and informs the other two, also providing them with information for updating their network topology models on which their diagnosis is based. The AAA starts to identify the specific network elements at fault - initially producing a quick, approximate answer which it sub-

1. "Towards a Cooperation Knowledge Level for Collaborative Problem Solving", N.Jennings, ECAI-92
2. "Using Joint Responsibility to Coordinate Collaborative Problem Solving in Dynamic Environments", N.Jennings & A.Mamdani, AAAI-92

sequently refines using a more accurate procedure. In parallel, the BAI starts determining the BOA, which when calculated is passed onto the AAA. In order to be consistent, the elements identified by the AAA should also be in the BOA produced by the BAI - a fact taken into account by the AAA during its detailed diagnosis. While the AAA and BAI are working on diagnosis, the CSI continues to monitor the network in order to detect significant changes in status or indicate whether the fault was only transient. Once a fault has been detected, each agent has a role to play and by combining their expertise, problem solving is enhanced. Overall system robustness can be improved by intelligently sharing information which is available in the system, but not readily available to all the agents. There are two main cases in which this can be seen: firstly if the CSI detects that the fault is transient, meaning the other two are attempting to diagnose a nonexistent fault. Secondly if further faults occur, the network topology may be so radically altered that the diagnosis is predicated on invalid assumptions. The role of joint responsibility is to provide the basis for determining which information should be shared and how agents should act when they receive it.

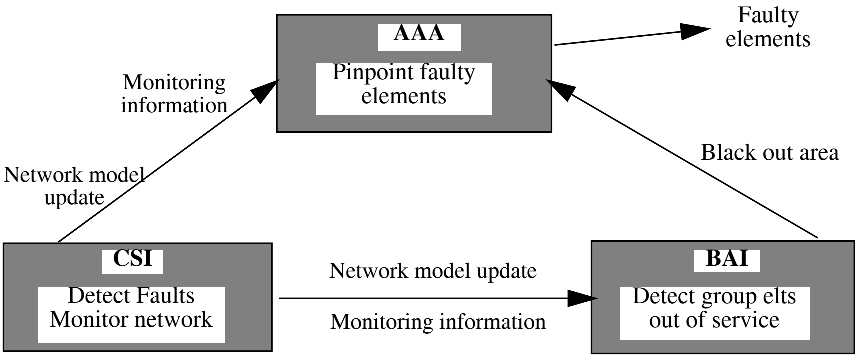

Figure 1: Cooperating Agents

3. Joint Responsibility

Joint responsibility defines preconditions which must be satisfied before joint problem solving can commence and prescribes how individual team members should behave once it has started.

3.1 Joint Problem Solving Pre-Conditions

Once the need for joint action has been established, three conditions need to be met before it can actually begin. Firstly, a group of agents who wish to solve a common problem must be identified. In our example, willing participants are those which have the goal of participating in the detection of faulty network elements. Secondly, participants must agree that they will work together to achieve their common objective - in particular they must acknowledge the *principle* that a common solution is essential. Without acknowledging this, there can be no *intentional* joint action, only unintentional (accidental) interaction. The actual solution will only begin to be developed once all prerequisites have been satisfied. Finally agents must agree that they will obey a “code of conduct” to guide their actions and interactions whilst performing the joint activity. This code, specified below, ensures that the group operates in a coordinated and efficient manner and that it is robust in the face of change.

3.2 Prescription of Behaviour

The notion of *commitment* is central to the definition of joint responsibility and ensures that once agents agree they will perform an action they will endeavour to carry it out. Therefore once the common solution has been agreed, all participants should ensure that they reserve sufficient resources to carry out the actions in which they are involved. However because of the unpredictability and dynamics of the environment - events may occur which affect this commitment. For example new information may become available which invalidates previous assumptions or unexpected events may require urgent attention. In such circumstances, it would be irrational for an agent to remain committed to the previously agreed actions; so conditions for reneging need to be enumerated. There are two levels at which lack of commitment can occur: to the common objective (eg there is no longer a need to diagnose faults) or to the common solution. The following are reasons for dropping commitment to the common objective[3]:

• **the objective already holds**; eg another agent has computed the faulty elements

• **the motivation for the objective is no longer present**; eg CSI realises that the group of alarms do not correspond to a fault

• **the objective will never be attained**; eg AAA realises that it is not being supplied with sufficient alarm messages to make a diagnosis

However conditions under which agents can drop commitment to the common solution also need to be defined. Separate conditions relating to plan states are necessary because dropping commitment to a plan typically involves developing a new solution for the same problem rather than dropping the goal completely (i.e. it has a different functional role) and also that it provides a more detailed specification for the system implementor. Reasons include:

• **following the agreed plan does not lead to the desired outcome**; eg CSI detects a substantial change in the network, meaning that the models being used by the AAA and BAI are so inaccurate that any ensuing diagnosis will be incorrect

• **one (or more) of the actions cannot be executed**; eg CSI is no longer receiving information about the network and so is unable to monitor its status

• **one of the agreed actions has not been performed correctly**; eg the BAI has been distracted by an unplanned task and cannot produce the black out area at the agreed time. Meaning the AAA cannot compare its initial hypotheses with the black out area to ensure consistency before undertaking the detailed analysis.

When an individual becomes uncommitted (to either the objective or the means of attaining it) it cannot simply stop its own activity and disregard other team members. Rather it must endeavour to inform all team members of this fact and also of the reason for the change. This ensures team members can monitor the progress of events which affect their joint work and, in the case of failure, the amount of wasted resource can be minimised.

As this scenario illustrates, collaborative activity is fraught with opportunities for inconsistencies and when it does run into problems, it is usually detected by only one team member. Without a prescription of how to behave or criteria against which to evaluate joint activity, the team may perform in an uncoordinated manner. The model of joint responsibility sketched here addresses both of these problems; defining conditions which are important for an agent to detect and prescribing actions which need to be taken in such circumstances in order to ensure maximum coherency in the group.

3. "Teamwork", P.R.Cohen & H.J.Levesque, SRI Technical Note 504