Controlling Cooperative Problem Solving using Joint Intentions

Nick Jennings
Department of Electronic Engineering
Queen Mary & Westfield College
University of London
N.R.Jennings@qmw.ac.uk

In Distributed AI (DAI) systems, problem solving agents cooperate to achieve the goals of the individuals and of the system as a whole. Such joint work is needed because of the dependencies between agents' actions, the necessity to meet global constraints and because often no one individual has sufficient competence to solve the entire problem. Initially I concentrated on developing a general-purpose cooperation framework (called GRATE) and applying it to the real-world industrial applications of electricity transportation management and diagnosis of faults in a particle accelerator controller. These experiments were successful in that it was possible to instantiate useful cooperation schemes, however when anything unexpected happened (eg new information invalidated existing goals, synchronisation between actions was disrupted or agents had misinterpreted the situation) the cooperating community became uncoordinated. This incoherence occurred because GRATE agents did not embody sufficient knowledge about the process of cooperative problem solving. To rectify this problem, it was decided that agents should have a well-grounded and explicit model of joint problem solving on which their behaviour could be based.

Existing models of joint intentions provide only a partial description of the process of collaboration. Most importantly, from the perspective of industrial applications, they fail to describe how joint actions may falter and how individuals and the group should behave in such circumstances. Also as the existing models were predominantly theoretical, little consideration had been given to computational tractability. To rectify these shortcomings, the model of Joint Responsibility was developed. This model specifies that each individual within a team should remain committed to achieving the common objective by the commonly agreed solution until one of the following becomes true: the objective has been met, the objective will never be met, the motivation for the action is no longer present, the desired outcome of a plan step is already available, following the agreed action sequence does not achieve the desired outcome, one of the specified actions cannot be carried out or one of the agreed actions has not been carried out. Whilst in this state the agent will honour its commitments to its agreed actions. However if an agent is no longer committed to the joint action or the common solution, it cannot simply abandon the action because its accomplices may not have been able to detect that there is a problem. For this reason, the responsibility model stipulates that when a team member is no longer jointly committed to the joint action it must ensure that all its acquaintances are informed of this change of state. This enables the whole team to reassess the viability of the joint action and in particular the actions involving the agent which is no longer committed.

A rule-based interpretation of joint responsibility was then used to build agents which had an explicit and principled model of collaboration to guide their individual and social actions. A series of comparative experiments were undertaken to assess the performance characteristics of these agents. Three types of problem solving organisation were compared - a responsible community; an implicit group model in which agents had individual intentions, but did not form explicit collaborating groups and groups of selfish problem solvers who set up joint intentions, but when the joint action became unsustainable, behaved selfishly and simply abandoned their local processing without informing their fellow team members. These experiments showed that

responsible communities performed more coherently than the other two; this difference being especially noticeable as the domain became more dynamic and unpredictable (i.e. the chance of joint action unsustainability increased). This gain in performance was achieved with negligible extra processing requirements for the coordination mechanisms.

This work shows, through empirical evaluation on a real-world problem, that a suitably formulated model of joint intentions is a powerful mechanism for coordinating the behaviour of collaborating agents. This is especially true in situations where agents have to make decisions using partial and imprecise data and when the environment itself is evolving and unpredictable. It also indicates how theoretical models of collaboration can be used as a basis for implementation level systems. Finally as a consequence of the insights gained in this work, a proposal for the next generation of multi-agent systems is made. In such "cooperation knowledge level systems" individuals maintain and reason about explicit and deep representations of social interactions, rather than having an implicit and shallow understanding of these processes.