Robustness for MultiAgent Systems in Hostile Environments

Alan White *(a.g.white@sms.ed.ac.uk)*

Supervisors: Austin Tate, Stephen Potter

Abstract: A MultiAgent System (MAS) is composed of multiple interacting agents, where the system purpose is met through achievement of individual agent goals. In a hostile environment, agents are subject to debilitation; resultant inability to achieve goals may have wide ranging consequences across the system. Approaches such as plan repair or conditional planning are invoked when a goal has been selected and failure has occurred; addressing the consequences of debilitation rather than the root causes of why actions fail to achieve the desired result. This work proposes a model relating capabilities to potential goals, including those from agent relationships. It is intended to improve robustness by allowing specification of pre-emptive, maintenance behaviour to prevent current debilitations from being allowed to cause future problems in MAS operations.

Proposed agent model

An agent *a* is modeled based upon the BDI approach[RAO99] for representing rational agents;

`a = <B, C, D, I, P>`

Where *B* is the set of *Beliefs* the agent holds about the environment (including knowledge of other agents capabilities), *D* is a set of potentially conflicting *Desires* (the set of potential goals), and *I* is the set of *Intentions* (concrete goals from *D*) the agent has decided to achieve.

The set *P* represents the *Policies* of the agent. This models how the agent responds to requests to perform tasks; specifically what types of *obligation* the agent will accept regarding that capability.

The *Capabilities* an agent holds are represented in the model in the set *C*. These may be primitive or compound; in the latter case involving structured actions, potentially including the ordered use of capabilities held by other agents.

Each $c \in C$ describes an action the agent can perform;

`c = <s, pre, eff, cs>`

Where *s* is a *signature* in the form $n(v_1, v_2, ..v_k)$ of name *n* and *k* required variables *v*. *Pre* and *Eff* respectively represent the state preconditions and effects of performing the action.

Finally, cs represents the *confidence of success* for performing the action, given a set of beliefs *B'*;

`cs(B') => [0...1]`

Where 1 indicates a 100% success rate in executing the relevant action (and thus achieving the resulting preconditions), in an environment where *B'* holds. This value varies dynamically; for example, debilitation to the agent, increased general workload or changes in the environment may result in a different probability the action would succeed.

There is a relationship between capability, desire and intention. Capabilities describe *all* the tasks an agent can perform, in unground form. Desires define a subset of more specific tasks as partially ground capabilities. Finally, the Intentions of an agent are fully ground desires, selected by meta-reasoning based upon beliefs, and met through the generation and execution of plans.

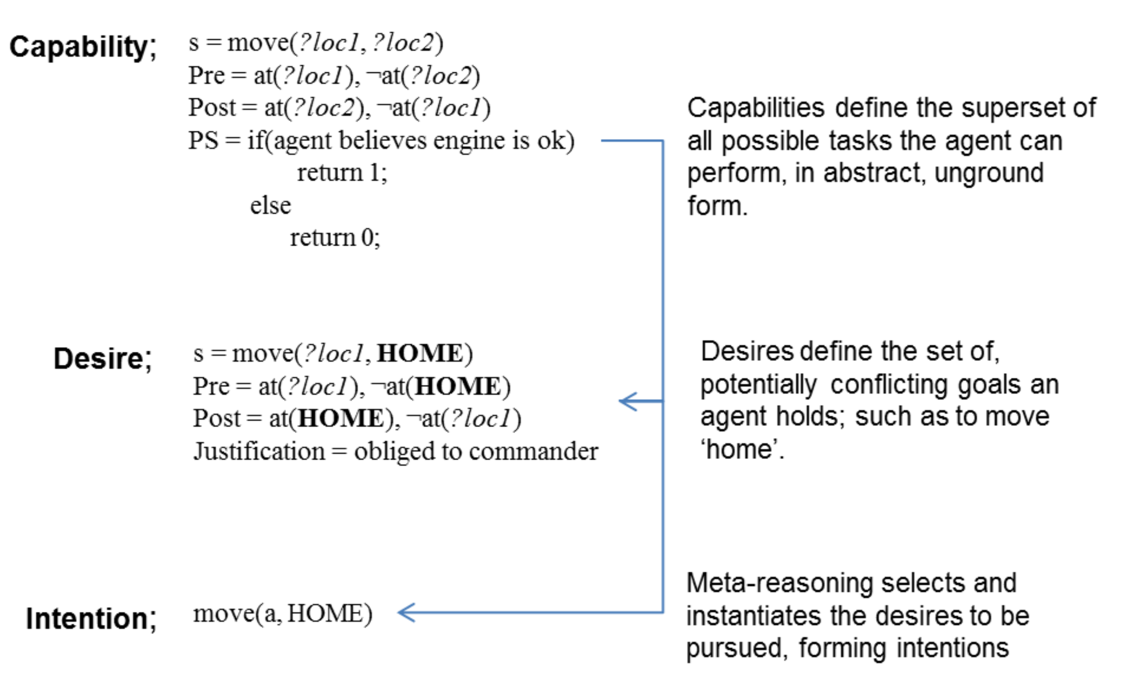

Fig 1; simplified example of how an agents 'move' capability relates to increasingly specified agent desires and intentions. A simplified function is used for to represent determination of the *ps* value.

Obligation / relationship formation

A key aspect of MASs is the formation of agent relationships to perform tasks together. An *obligation* is the responsibility an agent holds to perform a task for some other (dependent) agent - for example, as a result of planned task decomposition.

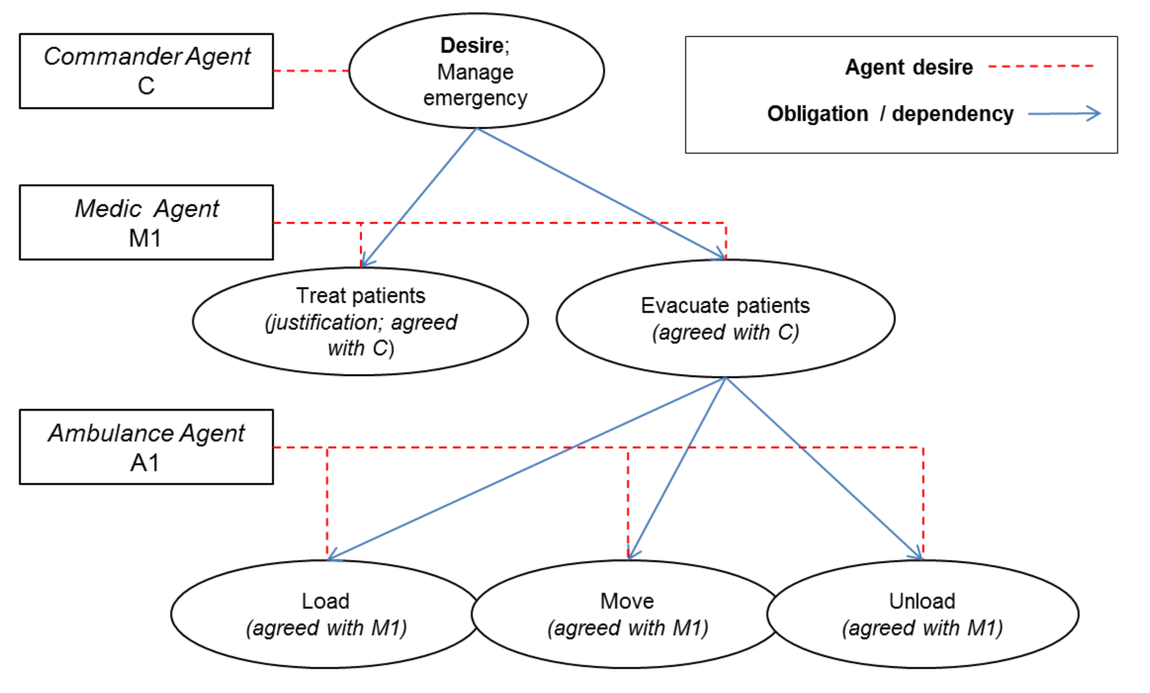

Fig 2; Example of obligations formed through action-plan decomposition; M1 holds a desire to be able to evacuate patients. This results in obligations upon an ambulance (transporting) agent to provide required capabilities for M1; to maintain achievability of the relevant desire.

Obligations are expressed through a *justification*, attached to desires held by an agent. The justification is used to describe when and why a desire is considered for selection as an intention, including whether or not the desire is held through an agents role in a dependency-obligation relationship.

An agent will accept obligations on the basis of a *policy*. The policy models how an agent will respond if there is a problem with meeting an obligation; what repair steps the agent will consider (if any) and the threshold conditions for these.

This includes an agent explicitly stating it will accept an obligation, but that it *cannot guarantee* it will constantly be capable of intending and meeting the associated desire.

Supporting robustness

Robustness can be defined as the maintenance of 'safety responsibilities'[WOOLDRIDGE99]; a safety responsibility is the prevention of some undesirable condition occurring.

In this context, these responsibilities are to ensure the achievement of intentions, when selected. Consequently, a robust system will be one that ensures, when a specific intention is selected, the agent has the capabilities to meet that goal state.

The intention is that agents will maintain task performing relationships, through the justification and obligation notions, constantly in order to maintain a 'state of readiness' for meeting future intentions. Intention selection meta-reasoning and the desire justifications can be used to predict the likely state where an intention will be formed; and thus assess capability performance under that state.

The policy under which an agent accepts an obligation guides both its own response to debilitation and assists the dependents consideration of contingencies or alternatives.

Potential reactions to debilitation

- **Drop obligation**; the obliged agent simply notifies the dependent it will no longer hold the desire (for example, if dependent has formed the obligation as a redundant backup)
- **Delegation**; the obliged agent assesses the debilitation as temporary in nature and identifies an alternate agent for the debilitative period (for example, due to temporarily heavy agent workload)
- **Replanning**; the capability is debilitated through some dependency on another agent or capability; the agent uses a planning type approach to compose a new approach using an alternate set of dependencies (for example, forming a new team).

Suggested evaluation domains

This work is aimed at domains where agent operate in a hostile, continuous environment, such that debilitation is a likely risk (with elements of unpredictability—i.e. a stochastic environment) and that there are time constraints upon reactive repair approaches.

One domain being considered with regards to these requirements is Pacifica[TATE94]; this is a fictional domain developed as a realistic model for studying Non-combatant Evacuation Operation (NEO) scenarios.

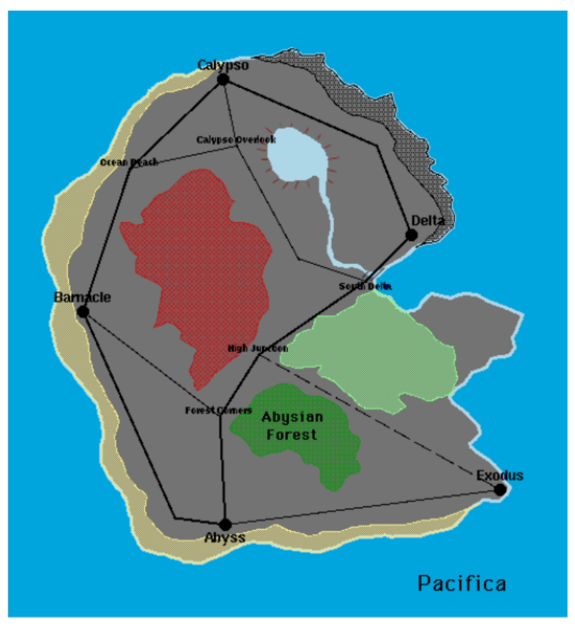

Fig 3; Geographical organization of the Pacifica domain

Operations in this domain have to account for physical geography and the impact of climatic events upon agents.

city	*Population centres*
road	*Point to point connections*
air-route	*Flight paths for aerial units*
bridge	*Connect roads over water*
airport	*Defines where aircraft can land*
seaport	*Where seaborne evacuation can be performed*
terrorist-group	*Oppositional forces representing a potential threat to agent operations*
ground-transport	*Road-constricted evacuation or cargo transport*
air-cargo-transport	*Cargo aircraft relevant to logistics tasks*
air-passenger-transport	*Passenger aircraft for evacuation purposes*

Fig 4; entity classes in the Pacifica domain;

Transport entities have range, fuel capacity, speed, etc. attributes. Terrorist-group entities represent a additional source of hostility in the environment.

Pacifica is being considered as it explicitly offers an environment intended for studying and evaluating failure management, in situations where co-ordination and time management is a key factor in success or failure of goals.

Progress and future work

Current work has revolved around the establishment of a formal agent model, to be used in the creation of reusable, generalized algorithms for agent operations and maintenance.

Agent robustness will be evaluated through simulated scenarios of debilitation, and the approach compared to alternative, reactive forms of repair. Suitable metrics will be specified, based upon the success of agents achieving their goals within specified time constraints.

It is argued that this agent modeling approach will lead to a system with intelligent maintenance behavior; improving robustness through addressing debilitation before it has the opportunity to impact intention formation and execution.

References

[RAO99] A. P. Rao and M. P. Georgeff, 1995. **BDI Agents – from Theory to Practice.** *In Proceedings of the First Intl. Conference on Multiagent Systems.*

[WOOLDRIDGE99] M. Wooldridge, N. Jennings & D. Kinney, 1999. **A Methodology for Agent-Orientated Analysis and Design.** *Proceedings of the 3rd International Conference on Autonomous Agents, AA99*. pp. 69-77.

[TATE94] A. Tate 1994. **PRECiS/Pacifica Suite of Scenarios.** *http://www.aiai.ed.ac.uk/oplan/pacifica/*

[KOMENDA08] A. Komenda, M. Pechoucek, J. Biba and J. Vokrinek, 2008. **Planning and Re-planning in Multi-actors Scenarios by means of Social Commitments.** In *Proceedings of the International Multiconference on Computer Science and Information Technology (IMCSIT/ABC 2008)*. IEEE, 2008. ISBN 978-83-60810-1.

THE UNIVERSITY of EDINBURGH informatics

EADS

Work funded by EADS Innovation Works