Governing Intelligent Virtual Agent Behaviour with Norms (Extended Abstract)

Jeehang Lee, Tingting Li, Marina De Vos, Julian Padget
Department of Computer Science
University of Bath
Bath, BA2 7AY, United Kingdom
{j.lee, t.li, m.d.vos, j.a.padget}@bath.ac.uk

ABSTRACT

One requirement by which virtual environments (VEs) are judged, is the believability of the virtual agents (VAs). One aspect of believability, is that agent responses to situations should not create cognitive dissonance and thereby distract the observer. One approach to this problem is the use of institutional models providing social reasoning, in conjunction with classical AI techniques providing individual reasoning, to achieve the appropriate recognition of complex situations and provide guidance on the subsequent choice of action(s). We present a distributed approach that offers governance – rather than regimentation – of intelligent virtual agents (IVAs) situated in a VE. We aim to show that the combination of an institution providing social reasoning and BDI agents providing individual reasoning, establishes a framework for enhancing believability through the interplay between: (i) the institution and IVAs in VEs, and (ii) norms maintained by the institution and the mental states of IVAs. From an engineering point of view, the framework provides a separation of concerns because the BDI agent is augmented with the capacity to process social obligations, while the specification and verification of social structure resides in the institutional models.

Categories and Subject Descriptors

I.2.11 [**Distributed Artificial Intelligence**]: Intelligent agents; I.3.7 [**Three-Dimensional Graphics and Realism**]: Virtual reality

General Terms

Design, Theory, Human Factors

Keywords

Intelligent Virtual Agents, Institutions, BDI Agent, IVA architecture

1. VIRTUAL AGENTS AND INSTITUTIONS

Our aim is to make avatars' reactions to social situations more believable. Our approach is separate from the reasoning about actions into two ways: (i) individual reasoning: provided by a BDI agent implemented on the Jason platform [2], which receives sensory data from the avatar and is ultimately responsible for communicating an action to the avatar, and (ii) social reasoning: provided by an institutional model – or as many as there are situations to model – implemented in Inst*AL* [3] and Answer Set Programming (ASP), which monitors the actions of all the agents in a given setting and advises each agent what is the socially correct action to take. As with any advice, the agent is free to choose to ignore it, but may be punished in some way consequently.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.
Copyright © 2013, International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

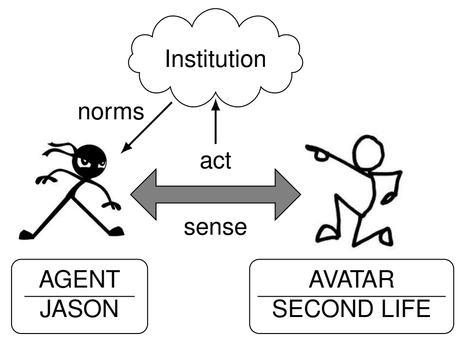

Figure 1: Architectural sketch

Architecture

We use a lightweight distributed framework introduced in [4], to provide communication, but also equally importantly, decoupling. The basic metaphor is of publish and subscribe streams, in this case implemented by an XMPP server. This permits the connection of multiple software components, either on a single computer, or several over a local or a wide area network. The connection between the agent and the avatar, is provided by the Openmetaverse [5] interface to the Second Life server: this supports not only the creation of VAs, but also the scripting of behaviours using combinations of various atomic VE actions. The framework is not tied to any of these components: different VEs can be used, as can different devices for connection to the real world.

The IVA is made up of two components: (i) the virtual character, that resides in the VE and being capable of sensing and acting in that environment, and (ii) the intelligent (BDI) agent, that resides on the agent platform, which senses outputs from the virtual character and generates actions that are inputs for the virtual character. The sensory data from the virtual character may either be raw or processed in some way. Likewise, the actions from intelligent agent may either be precise actions or high level commands to be broken down. Thus, decision making is carried out by the BDI agents, and each BDI agent is directly mapped to a virtual character. The beliefs of the BDI agent are established by percepts received from the VA, along with norms (permissions, obligations), received from the institution(s). The role of the institution is in formulating and communicating norms corresponding to environmental events delivered by BDI agents. These norms may then become a part of a belief set in the BDI agent and hence part of the agent's decision-making process.

Mental Model

The mental model of the distributed agent framework is somewhat different from a conventional agent platform, because changes of mental state not only depend upon percepts from the VAs, but also on the norms from the institution that are incorporated as percepts. Given an established VE, the actions of the IVAs bring about observable changes in the environment, which are perceived by the virtual characters and delivered to both the BDI agents and the institution as events ($\mathcal{E}$). Thus, the virtual characters act as sensors and filters for both BDI agent and institution and in consequence the agent and the institution are both aware of the same VE events, which we call *external events* ($\mathcal{E}_{ex}$), because they occur external to these entities. We use event to refer to the symbolic representation of some captured data about the environment. The *Recognising* component in the VA is responsible for turning whatever is observable in the VE into this symbolic representation. We assume that the perception and recognition of all $\mathcal{E}_{ex}$ is carried out by the VA. All such events are delivered to the BDI agent, but only some may be meaningful to the agent, and those that are not are ignored. The same applies to the institution: if there is no constitutive rule for an $\mathcal{E}_{ex}$, then it is not of relevance for the institution. It should be noted that an event that is not meaningful in itself to the agent, but is to the institution, may consequently bring about an institutional change that is in turn relevant to the agent, since the event may cause the institutional recognition of a situation that affects the normative position of the agent. Thus, the institution consumes $\mathcal{E}_{ex}$ from the VE and provides a social interpretation of them to the co-located IVAs, in terms of normative information (permission, obligation), which is then incorporated into their decision-making process. In so doing, the agents may thus exhibit socially aware responses to situations, which may be perceived by humans as more believable.

Interplay of Institution and IVA

The institution provides two services for the BDI agent: (i) the social interpretation of VE events and delivery of normative position updates, and (ii) a query mechanism, whereby the BDI agent can ask the institution for information about the normative position.

The sequence of the runtime reasoning model [1] is described as follows. (i) The VA perceives information about the environment (ii) *Percepts* are delivered to the BDI agent which updates its beliefs (iii) *Events* are delivered to the Institution, which updates the normative state (iv) *Queries* are delivered to the Institution, which triggers a reasoning process (in ASP) and replies with information about the normative state, which the agent updates its beliefs (v) *Normative information* is delivered to each BDI Agent, which updates its beliefs. Thus, the IVA is able to extend its belief base using the information acquired via the social reasoning capability of the institution.

Normative information and IVA Mental States

In the mental model, normative information is complementary information for each agent rather than a separate part of the mental state. If an institutionally generated obligation were automatically adopted as a final goal, the agent would in effect be regimented, thus the model presented is capable of dropping back to fully regimented behaviour, if the agents have the appropriate behaviours. The actual form of the normative information is represented as: $\mathtt{obl}(\mathtt{act}, \mathtt{deadline}, \mathtt{violation})$ or $\mathtt{perm}(\mathtt{act})$, which means that an agent *X* is obliged to perform action *act*, or an agent *X* is permitted to perform action *act*, respectively. A BDI agent will then take action *act*, if it satisfies a belief or a subgoal of its main goal. If not, the obligation is ignored and a sanction may follow.

Illustrative Examples

Queueing: It is common in the society modelled in this scenario, that anyone who wishes to enter somewhere, or get on something is obliged to wait their turn in a queue. Individuals normally wait in order of arrival, but if someone arrives who is disabled or old, the social obligation is for those queuing to make a space for them at the front.

Inter-Personal Distance: Inter-Personal Distance(IPD), or proxemics, is the personal space between an individual and others. It has been observed that human-controlled avatars appear to be subject to the same IPD social norms as in the real world. Therefore, VA behaviour may be more believable is such norms are followed.

The social norm in relation to intimacy and IPD is: if people get physically too close to a person with whom they are not sufficiently intimate, then they typically either change their eye gaze or move to keep the proper level of IPD.

2. CONCLUSIONS AND FUTURE WORK

We have set out a framework for more believable IVAs, that works by providing better (social) situational awareness, while avoiding burdening the agents themselves with substantially more knowledge or rules about every situation they might encounter. This is achieved by a flexible, distributed communications middleware that enables the connection of VAs in a VE, BDI agents and a regulatory institutional model.

The separation of institutional knowledge should help with authoring, validation, deployment and maintenance. A challenge for future work are the problems arising from institutional overlap – that is, when more than one institution recognises a situation. A further challenge is to support rule change over time and for new rules to be created and even whole new institutions. Both norm emergence, leading to the creation and incorporation of new rules and norm revision, would seem to be essential features for the long-term utilisation of institutions in VEs.

3. REFERENCES

[1] T. Balke, M. De Vos, J. Padget, and D. Traskas. On-line reasoning for institutionally-situated bdi agents. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3*, AAMAS '11, pages 1109–1110, Richland, SC, 2011. International Foundation for Autonomous Agents and Multiagent Systems.

[2] R.H. Bordini, M. Wooldridge, and J.F. Hübner. *Programming Multi-Agent Systems in AgentSpeak using Jason (Wiley Series in Agent Technology)*. John Wiley & Sons, 2007.

[3] O. Cliffe, M. De Vos, and J. Padget. Specifying and reasoning about multiple institutions. In *AAMAS06 Workshop on Coordination, Organization, Institutions and Norms in agent systems (COIN-2006)*, 2006.

[4] J. Lee, V. Baines, and J Padget. Decoupling cognitive agents and virtual environments. In Frank Dignum, Cyril Brom, Koen Hindriks, Martin Beer, and Deborah Richards, editors, *Cognitive Agents for Virtual Environments*, volume 7764 of *LNCS*, pages 17–36. Springer, 2013.

[5] OpenMetaverse Organization. libopenmetaverse developer wiki. http://lib.openmetaverse.org/wiki/.