Toward Societal Acceptance of Artificial Beings

Daniel I. Thomas
Technology One Corp,, Australia

Ljubo B. Vlacic
Griffith University, Australia

INTRODUCTION

Modern organizations are faced with many challenges with the trend toward distribution of their workforce across the planet. With this situation becoming more common, it is important for organizations to find ways of encouraging effective leadership and strong teamwork. Training and evaluation of the effectiveness of those employed can be an expensive exercise due to geographic separation of the parties involved. To this end, we propose a collaborative play scenario, using humans and artificial beings as *fully equal partners* (FEPs), to facilitate training and evaluation of a dispersed workforce.

While this scenario is a simple example of collaboration among human and artificial entities, moving this concept forward in other application areas creates questions about how artificial entities influence outcomes in the context of group decision making. The idea of social influence and acceptance of artificial beings as equal decision makers is explored, and how they may integrate into larger societies.

In this article, we present a simple training exercise designed to test a candidate's leadership ability to negotiate with other members of the organization, using their influence to achieve (partially or fully) their goals. While in practice, the play scenario would consist of combinations of human and artificial beings, the training scenario presented shall consist solely of artificial FEPs in order to demonstrate how influence can affect a result in a collaborative process.

BACKGROUND

Artificial Beings as Fully Equal Partners

When the average person is confronted with the term "Artificial Intelligence" it is more likely to conjure images of science fiction than science fact (Khan, 1998). Yet throughout our daily lives, we experience various degrees of artificial intelligence in such mundane devices as washing machines and refrigerators. Beyond this, organisations have been using intelligent systems in a myriad of endeavours. Beyond today however, "We may hope that machines will eventually compete with men in all purely intellectual fields" (Turing, 1950, p 460). Turing's remarks may not be fully realized today; however, the integration of artificial beings into human organizations and society evoke powerful images of both positive and negative possibility.

One possibility is artificial beings emerging as partners rather than tools in various collaborative situations. Unlike past revolutions of mechanical automation, the presence of artificial beings should not imply a redundancy for human partners, but rather a complimentary relationship. Group decision making, including both humans and artificial beings as equals, increases the diversity of the knowledge pool (Dunbar 1995), improving the likelihood of positive outcomes.

In order for artificial beings to be realized as collaborative partners, as opposed to an intelligent tool, they must be able to articulate their perspectives and opinions, while taking onboard the knowledge and opinions of others. For this to occur, artificial beings require a degree of social influence. For this influence to occur, the artificial being needs to become acceptable within the social system: Society, organization or group (Kelman, Fiske, Kazdin, & Schacter, 2006). In making the transition to societal acceptance of artificial beings, there are great challenges, both technical and social. To better study artificial beings as collaborative partners, it is possible to focus on a smaller, group social setting, with an assumption of social acceptance (and therefore the capability to influence) collaborative group decision making. For this reason, computer games provide an excellent environment for understanding how humans and artificial beings can positively influence outcomes in a collaborative group situation.

Much of our work into collaboration has been influenced by the use of intelligent autonomous agents in computer games. Jennings and Wooldrige (1995) describe an intelligent agent as one that enjoys the attributes of autonomy, situatedness, social ability, reactivity and proactiveness.

Basing intelligent entities around this core concept of agency has led researchers such as Laird (2001) and Kaminka et al. (2002) to create intelligent opponents for human players.

Taking this a step further, we see future applications for intelligent artificial beings as more than just opponents or nonplayer characters (called NPCs) in computer games, but rather we see artificial beings being utilised as *fully equal partners*.

Extending these concepts of humans and artificial entities interacting collaboratively in computer games, it is necessary to define a type of entity that:

- Does not treat human and artificial players differently during interaction;
- Can work cooperatively with other fully equal partners (including humans);
- "Plays" the game as a human would;
- Does not work to a defined script or take direction from an agent "director" such as those described by Magerko et al. (2004) and Riedl, Saretto, and Young (2003) and;
- Is not necessarily aware of the nature of other FEP beings (human or artificial in nature).

Simply, a *Fully Equal Partner* (or *FEP*) is an intelligent entity that performs tasks cooperatively with other FEPs (human or artificial), but is also capable of being replaced one with another. These beings are not necessarily aware of the nature of their fellow partners.

A Collaborative Architecture

In order to facilitate the collaboration among fully equal partners, the involved computer games must support a number of key features (Thomas & Vlacic, 2003), including:

i) A clean and well-defined interface or separation between the beings and the game (Vincent et al., 1999);
ii) A concept of time and causality; and
iii) Support for experimentation (Cohen, Hanks, & Pollack, 1993).

To create a computer game that enjoys many of these features, Thomas and Vlacic (2005) developed a layered architectural approach to collaborative games. Collaborative computer games that involve FEPs have three architectural layers: A communications, a physical and a cognitive layer.

The communications layer is essentially the protocols and low-level software that facilitate interaction and communication. The physical layer describes items and entities within the game world and how they may be manipulated. The cognitive layer describes the processes required to facilitate intelligent collaboration.

Collaborative process

If considered at a high level, the collaborative process *c* involves taking a set of fully equal partners *P* with a set of goals *G* and producing a set of outcomes *O*. An outcome may not necessarily satisfy the set of goals (e.g., a failure outcome).

$$O = c(P,G)$$

In order to obtain these outcomes to the collaborative process, FEPs engage in conversations. The result of these conversations are pieces of group collective knowledge *K*; that is, knowledge that is known to the group. Outcomes of the collaborative group are a result of the collaborative process between the group of FEPs and the goals of the collaborative process.

$$O = c(P,G)$$
$$O = c(P,G)$$
$$O = \{o_1,..,o_n\}$$
$$o_n = s(P, n(G, K^P))$$

where *s* is a function of all partners *P* applied to an interpretation function *n* of the set of goals *G*, the set of group collective knowledge across the entire set of partners K^P, resulting in an outcome o_n.

INFLUENCE IN THE COLLABORATIVE PROCESS

In human to human interactions, we see many forces at play that influence one person to agree or take the side of another in a discussion. These influences need to be taken into account when collaborative work is undertaken. Even the size of a group (Fay, Garrod, & Carletta, 2000) can change the way in which partners are influenced, and by whom.

Collaborative FEPs may create an affinity with one or more entities and are more likely to accept their position during negotiation. Possible methods for obtaining an affinity with one or more FEPs include:

1. The degree to which one FEP's responses convey a perception/opinion that matches that of another FEP. The more that one partner's position matches that of another partner, it becomes more likely that the partner will "trust" the statements of that partner.
2. Some arbitrary/authoritative influence factor that has the partner tending toward the position of one or more