2011

Representing Trust in cognitive social simulations

Pollock, Shawn S.

Proceedings of the 4th International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction (SBP'11), 2011, pp. 301-308.
http://hdl.handle.net/10945/36617

Representing Trust in Cognitive Social Simulations

Shawn S. Pollock, Jonathan K. Alt, and Christian J. Darken

MOVES Institute, Naval Postgraduate School
Monterey, CA 93943
{sspolloc,jkalt,cjdarken}@nps.edu

Abstract. Trust plays a critical role in communications, strength of relationships, and information processing at the individual and group level. Cognitive social simulations show promise in providing an experimental platform for the examination of social phenomena such as trust formation. This paper describes the initial attempts at representation of trust in a cognitive social simulation using reinforcement learning algorithms centered around a cooperative Public Commodity game within a dynamic social network.

Keywords: trust, cognition, society.

1 Introduction

One of the fundamental phenomena governing human interactions is the notion of trust, without which the fabric of society quickly comes unraveled. Trust in other humans and societal institutions facilitate a market economy and a democratic form of government. The human information processing system, adept at receiving and synthesizing large amounts of sensory information from our environment, manages to identify which percepts are salient to our current task or our long term well being by the allocation of selective attention [1]. In this view the level of trust is a quality associated with each percept either directly, one sees an event as a first person observer, or indirectly, the information comes from a reliable source. The latter case, involving the interaction and development of trust between individuals, will be the topic of this paper. The representation of trust within cognitive social simulations is of fundamental importance to the exploration of macro level social phenomena.

When humans interact with one another there is a multitude of information, both verbal and nonverbal, that is exchanged between the participants. In social simulations the major issue in modeling this phenomenon is to understand more fully how information flows through a social network and how the society develops and evolves its beliefs over time [2]. Central to the modeling of information flow in social systems is the concept of trust. The main contribution of this paper is a model of trust based on reinforcement learning and demonstrated in the context of the Public Commodity Game [3].

This paper first provides a brief introduction to trust, reinforcement learning, and cognitive social simulation. This is followed by a description of the general trust model as well as the current Python implementation of the model within the context

J. Salerno et al. (Eds.): SBP 2011, LNCS 6589, pp. 301–308, 2011.

of the Public Commodity game. Computational results of experimentation with model parameters related to the formation of norms and penalties for the violation of norms are provided as well as a summary and discussion of future work.

2 Background

This section provides an overview of the notion of trust as used in this setting, reinforcement learning, and cognitive social simulation.

2.1 Trust

In order to model trust, it is first necessary to define trust. Trust is not a simple concept (though it is easy to express its general intent) and codifying trust to the extent of incorporation into computer simulation is rather difficult. This work establishes a definition of trust suitable for use in simulation rather than a general definition. In particular the definition of trust used is chosen with an eye toward the long term goal of the research, to model communication and belief revision. Trust is viewed as an analysis performed by an individual agent to prejudice newly obtained information either based on the sender or the topic of discussion, and the history between the sender and the receiver and topic [4]. As an example, if a person who is completely trusted by another shares some information, the receiver is more likely to welcome the communication and take action on it. Additionally, if the sender is moderately trusted by the individual, but the receiver is untrusting of the topic (such as a particular political view), they may disregard the information as non-actionable. Simply making a binary decision of actionable versus non-actionable information is not robust enough to capture the nuances of human behavior requiring the incorporation of some concept of the grey area in between. The model incorporates sender and topic into the state space of a reinforcement learning algorithm to develop a two-pass notion of trust where the receiving agent first determines whether information is actionable then makes a separate decision if he should or should not revise his beliefs based on this new information.

2.2 Reinforcement Learning

An appealing approach to represent human like learning and action selection is the idea of reinforcement learning, where agents will seek select actions within their environment based on their experience. Based on the permissiveness of the environment, agents are eligible to receive percepts from the environment that inform them on the state of the environment at a given point in time. The basic elements of reinforcement learning are: a *policy* that maps states to actions; a *reward function* that maps a state of the environment to a reward; a *value function* that maps states to long term value given experience; and an optional *model* of the environment [5]. The policy provides a set of actions that are available in a given state of the environment; the agents leverage their prior knowledge of the environment, informed by the value function, to determine which action will provide the greatest reward, as defined by the modeler. Agents must strike a balance between exploration, behavior to explore the reward outcomes of state action pairs that have not been tried, and exploitation, behavior that takes advantage of prior knowledge to maximize short term rewards, in order to avoid converging to local minima [5]. The ability to control this balance makes reinforcement learning an attractive approach for representing human behavior. The reinforcement learning technique used in this work is Q-learning in conjunction with a softmax function (the Boltzmann distribution).

Q-learning, $Q(s,a) \leftarrow Q(s,a)+\alpha(r+\gamma max_{a'}Q(s',a')-Q(s,a))$, falls into a class of model free reinforcement learning methods that have the property that the learned action-value function, Q, approximates the optimal action-value function, Q*, requiring only that all state action pairs be updated as visited [5]. For each state action pair, *(s,a)*, the Q-learning function updates the current estimate based on new information received from recent actions, *r*, and discounted long term reward. In general, an action is selected from a given state, the reward outcome is observed and recorded, and the value function updated. The value associated with each action is used during each visit to a particular state to determine which action should be chosen using the Boltzmann distribution, shown below.

$$P_i = \frac{e^{\hat{Q}(\sigma,a)_i/\tau}}{\sum_j e^{\hat{Q}(\sigma,a)_j/\tau}} \tag{1.1}$$

The Boltzmann distribution uses the temperature term, τ, to control the level of exploration and exploitation. A high temperature translates into exploratory behavior, a low temperature results in greedy behavior.

2.3 Cognitive Social Simulation

Cognitive social simulations are particularly well-suited for defense applications which typically call for analysis of a military force's performance while operating as part of a complex conflict ecosystem [6]. Agent based models have been applied to the military domain previously [7], but the use of cognitive social simulation, enabled by cognitive architectures, in this application area is relatively new [8]. The relevancy of these tools is particularly highlighted by the nature and objectives of the current conflicts, where the population of the conflict area is seen as the center of gravity of the operation [9]. Gaining an understanding of potential means of transitioning the social system in these countries from an unstable to a stable state provides a challenge to leaders at all levels of the military and civilian organizations involved. The representation of trust and its impact on the effectiveness of information operations is required in order to provide greater insight into the population within the area of interest.

3 Approach

This section provides an overview of a general model of learned trust for agents in cognitive social simulations, an introduction to the test bed environment, and a description of the proof of principle implementation in Python 2.6.

3.1 General Model

This general model is a turn-based simulation in which the agents and their relationships are represented by a single network graph with the agents as the nodes and their social relationships as the edges, weighted by the value of the relationship. To be specific the network graph is a combination of two similar graphs with identical nodes, but in one the edges are bidirectional and represent a base constant value and the second one in place of each bidirectional edge contains a pair of directed edges representing the agent's individual variable contribution to the strength in the relationship. In this way, the edge weights have both a static and dynamic component. The static is based entirely on the concept of homophily (E_H), that like persons associate more frequently, utilizing a simple Euclidean distance calculation and as stated above is set as a constant [10]. The dynamic portion is completely under the control of the agents involved ($E_{A\rightarrow B}$). In the case of k total agents, each agent has k-1 choices of who to spend time with and based on that emphasis will receive some unknown reward from these relationships. For every simulation round, each agent will choose to increase, decrease or maintain their contribution to their relationships with the other agents. This contribution can be seen like a fraction of time spent with the others in that it is represented as a floating point number from 0.0 to 1.0 and such that the sum of all these components (i.e. the sum of all edge weights leaving the node representing the agent) always sums to 1.0. At the end of each turn the agent is rewarded based on the strength of their relationships. This reward takes the following form:

$$Reward = E_H \cdot \min\,(E_{A\rightarrow B}, E_{B\rightarrow A}) \tag{1}$$

The equation uses the minimum of the variable contributions from each agent. In this way it more accurately can be said that the variable portion is the fraction of time that the agent would *like* to spend with the other agent and therefore using the min also provides a small penalty for those who place more emphasis on a relationship than the recipient.

The result of this basic model is the development of a simple dynamic social network. The network tends to become highly centralized around 1 or 2 agents. In particular, in runs consisting of 50 agents, the final network graph consisted of nearly every agent with a strong connection to a single central agent with no other connections present. In order to mitigate this affect it was necessary to add a second order factor in the reward calculation for the agents. Continuing along the analogy that the emphasis represents a fraction of time desired to be spent with the other agents, then it is natural to extend this and allow for groups of more than 2 agents to spend time together. In other words if two agents have a strong relationship and also each have a strong relationship to the same third agent, then all three agents should receive an additional reward for this.

The second order reward factors are based on the same reward function as used in the first order above. In this case, the reward is divided by a distribution factor and subsequently squared. For the case of agents A and B as above, but this time having a common friend in agent C, the additionally reward looks as below:

$$2nd - Reward = \min(\,E_H \cdot \min(E_{A\rightarrow C}, E_{C\rightarrow A})\,, E_H \cdot \min\,(E_{A\rightarrow C}, E_{C\rightarrow A}))\;/\;D \tag{2}$$

The closeness centrality of the network is highly sensitive to the distribution factor and is discussed in detail in section 4.

Once the second order terms are added similar network properties to what we would expect to see in real social situations emerge; namely subdivisions into clique's, pairings and the exclusion of certain individuals. It is obvious that this feature is not intended to actually model the internal and external processes that form human social networks; rather it is simply building the stage on which to build future work.

3.2 Prototype Trust Implementation

Each agent has a simple belief structure consisting of a finite set of beliefs, five in this case, represented by a single floating point number from 0.0 to 1.0.These beliefs combine in simple linear combinations to provide issue-stances, one in this case, also as a floating point from 0.0 to 1.0. During each turn of the simulation, following an initial stabilization period, set initially to 1000 rounds for a simulation of 15 agents, the agent will choose a topic based on a probabilistic Boltzmann distribution, and discuss this topic with its k nearest neighbors. Other than the k nearest there any neighbor above a specified threshold will always receive a communication, while neighbors below another threshold will never receive a communication.

Initially, the communications consist of each agent telling his closest neighbors what his value is on a selected belief. The receiving agents then use a reinforcement learning algorithm to determine whether or not belief revision is merited. In order to utilize reinforcement learning it is necessary to define some concept of a reward that the agent will receive based on their beliefs and therefore directly related to their trust and belief revision mechanisms. Our inspiration for a reward model comes from Game Theory and is called the Public Commodity Game [3]. In this game, each agent has an option to contribute to a public pot of money each round or to opt out. Following the round the money in the pot is multiplied by some amount (in this case 3.0) and then redistributed to each agent regardless of contribution. As a slight variation of this, the agents are allowed to decide an amount to contribute rather than to choose from full or no contribution.

In the current model, agents are given 1.0 possible units to play such that an agent that contributes nothing is guaranteed a reward of at least 1.0 for opting out and an unknown reward ranging from nearly 0 to 3.0 for full contribution. Game theory tells us that without cooperation the expected equilibrium for rational players would be exactly 0.0 contributions from all agents, in other words all agents take the guaranteed 1.0 and opt-out of the public commodity all together [3]. It is likely the case that some people would always contribute at least some small amount irrespective of their losses. In order to achieve this, "Faith in the Public Commodity" is the issue-stance and is used to directly control the level of their contribution to the public commodity. During each simulation round, agents communicate with one another and attempt to bring other agents closer to their beliefs. Despite the fact that a 1.0 contribution from all agents is the most mutually beneficial strategy, it is not a stable equilibrium as a single person could quickly realize that decreasing their contribution will increase their total revenue and would be easily reinforced by learning algorithms. What is seen is the agents benefit the most from a strategy of trying to make all the other agents contribute at a level consistent with their issue strength.

Now that there is a concrete idea of reward in this model, it is possible to begin applying a simple model of belief revision. Essentially, the agent will use a reinforcement learning algorithm where the state space is represented by sender-topic pairs and the action will be to dismiss the communication or to apply a revision, based on their level of trust in the information, resulting in an update of the fraction of time they desire to spend with the sender. Without any further weighting factors or limiting algorithms, the expected result of this simulation is that the beliefs of all agents will approach a stable equilibrium in which all beliefs are the same.

4 Experimentation and Analysis

This section provides a comparison of model results from several attempts to moderate belief revision within the model.

4.1 Experimentation of Second Order Factors in the Base Social Network

The first task was to evaluate the effects of the strength of the second order terms in the base social network reward functions. The distribution factor was varied and showed a fairly steep "S" curve that was centered between D = 14 to D = 24.

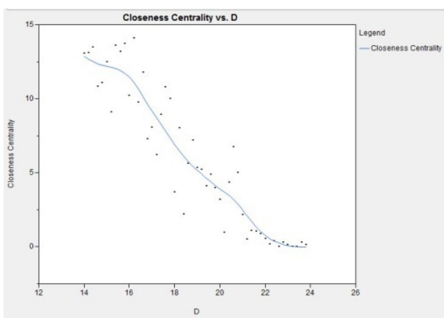

Fig. 1. Closeness centrality versus distribution factor

The distribution factor in affect allows fine tuning of closeness centrality in the base model in order to fit it to any particular purpose. There are several widely varying sources on what a real human social network should look like in terms of closeness centrality that range from 0.20 to 0.60. Therefore, for the purposes of the remainder of this initial experimentation D = 18.4 is used in order to target the fractional closeness centrality to around 0.30. The exact nature of these values is irrelevant for this initial model and only serves as a baseline for further work.

4.2 Belief Revision Experimentation

As discussed previously it is necessary to implement within the model a method for moderating how the agents can revise their beliefs. Initial methods for doing this are, at least for now, centered on a penalty for changing beliefs.

$$Net\ Dividend = \left(1.0 - Contribution\right) + \ 3.0\ x\ Contribution - NormPenalty$$
$$NormPenalty = \ e^{F*BeliefVariance}$$

Where the Belief Variance is a simple Euclidean distance measure from the agents current beliefs to what they started with at the beginning of the simulation. What is surprising is that when F is varied, there is no marked difference in the outcome of the simulation from a purely statistical point of view. What is seen however is in the Public Commodity play over time.

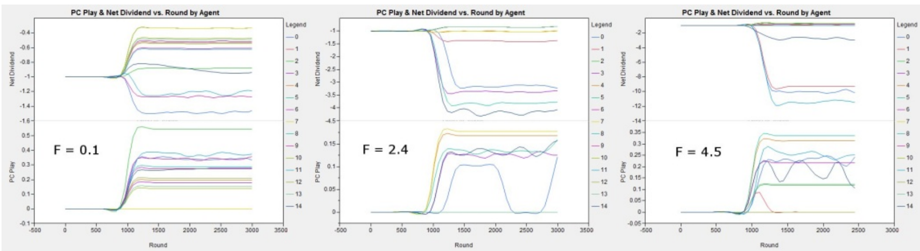

Fig. 2. Public commodity game play versus norm penalty

The higher the factor F becomes, the more unstable the Public Commodity game is. In other words, with a small norm penalty the agents will tend to find a stable equilibrium and remain there with fairly significant stability. As F is increased, the stability is decreased. The intriguing thing is that this behavior appears to be similar to actual human interaction. For example, if we look at a society there is a sense of a norm although it will really change over time it will remain fairly constant over small enough time periods. In this society there will be people or factions that challenge the social norm causing brief unstable equilibrium away from the norm that seem to return to the social norm after some time. More investigation into this phenomenon is currently underway. Once there is a satisfactory explanation for this behavior it may be possible, just as in the second order parameter, to tune the magnitude of the norm penalty in a way that is unique to the society being modeled.

5 Conclusions and Future Work

The results of the initial work are promising in that it is seen that there are factors within this simple first draft model that allow the social network to form and evolve based purely on the agents cognitive processes however the nature of the network can be tuned to match a desired case. This initial effort to model trust shows a lot of promise and merits further investigation. The next step in this process will be to utilize this algorithm within an existing social simulation and evaluating the effects.

The next generation of this model is planned to include a much more complicated belief structure that will include several metacognitive elements giving the agents some control over belief revision. There will also be included within the belief structure mental models of the other agents and the environment so that a perceived closeness in beliefs of the other agents will play directly into trust.

References

[1] Anderson, J.: Cognitive Psychology and Its Implications, 6th edn. Worth Publishers, New York (2005)
[2] Sun, R.: Cognitive social simulation incorporating cognitive architectures. IEEE Intelligent Systems, 33–39 (2007)
[3] Owen, G.: Game Theory. Academic Press, San Diego (1995)
[4] McKnight, D., Chervany, N.: The Meanings of Trust. University of Minnesota (1996)
[5] Sutton, R., Barto, A.: Reinforcement Learning: An Introduction. MIT Press, Cambridge (1998)
[6] Kilcullen, D.J.: Three Pillars of Counterinsurgency. In: US Government Counterinsurgency Conference
[7] Cioppa, T.M., Lucas, T.W., Sanchez, S.M.: Military applications of agent-based simulations. In: Proceedings of the 36th Conference on Winter Simulation, p. 180 (2004)
[8] National Research Council, Behavioral Modeling and Simulation: From Individuals to Societies. National Academies Press, Washington DC (2008)
[9] C. O. S. W. Department of the Army, Counterinsurgency. U.S. Army (2006)
[10] McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual Review of Sociology 27(1), 415–444 (2001)