Refining Human Behavior Models in a Context-based Architecture

David Aihe and Avelino J. Gonzalez

Intelligent Systems Laboratory
School of EE and CS
University of Central Florida
Orlando, FL 32816-2450
aihe@aihe.com
gonzalez@ucf.edu

Abstract

This paper describes an investigation into the refinement of context-based human behavior models through the use of experiential learning. Specifically, a tactical agent was endowed with a context-based control model developed through other means and tasked with a mission in a simulation. This simulation-based mission was employed to expose the agent to situations possibly not considered in the model's original construction. Reinforcement learning was used to evaluate and refine the performance of this agent to improve its effectiveness and generality.

Introduction and Background

How one makes a decision when faced with a task can be described as that person's behaviour. The Oxford dictionary [1], defines "behaviour" as 'the actions or reactions of a person or animal in response to external or internal stimuli'. Human behaviours are, consequently the actions or reactions of a human in response to some external or internal stimuli. The external stimuli include touch, smell, sight, and others. How a person reacts to these stimuli dictates his or her actions at that point in time.

Many factors affect how a human behaves in a given situation. For example, one would expect a person being held at gunpoint to cooperate with his or her captors. However, could this behavior in all certainty, represent every person? What if a martial arts specialist sees an opportunity to overcome his/her captors? Would this person react similarly? This suggests that there are multiple ways humans behave in particular situations. The number of variables involved in a person's action in a given situation is large and, as such, trying to address these variables would lead to unsolvable problems or representations that don't adequately fit the situation. With this in mind, researchers [2, 3] have proposed representations to address very specific situations. Most of these situations are referred to as *tactical situations* and have a smaller number of variables, though this limited number of variables could also lead to having a complex problem.

Gonzalez & Ahlers [3] propose and implement a paradigm called *Context-Based Reasoning* (CxBR) that can model a humans' expected behavior in any situation. CxBR models human behaviors in terms of contexts. This method seeks to limit and reduce the complexity inherent in human decision-making by limiting the number of events available for the agent to consider in any given situation. Several successes have been achieved using this method, for example [4, 5] among others. However, although an effective method, models are not always developed considering all possible situations experienced, or potentially experienced by an expert.

This research describes a method that addresses the refinement of context-based models of human behavior. To achieve this, *Reinforcement learning* (RL) is synergistically incorporated within CxBR. Reinforcement learning is a machine learning strategy that assigns rewards (positive or negative) as an agent (simulated or live) interacts with its environment (immediate or distant). The synergistic combination of these methodologies promises to significantly enhance the ability of CxBR to represent human tactical behavior.

Conceptual Approach

The approach used to address the problems described above utilizes experiential learning (reinforcement learning). This approach is based on the popular saying "20/20 hindsight". The human behavior model created from the acquired knowledge is subjected to real-life scenarios in a simulator. The agent constantly interacts with its environment in a simulator and adjusts its future actions (foresight) based on its past actions and rewards / punishments received (hindsight).

This research is not the first attempt at synergistically combining reinforcement learning and contexts. Wan & Braspenning [6] propose an extension to the RL framework to incorporate the role of contexts in solving RL problems. They had encouraging results in an experiment where the agent had to learn to intercept a moving target from any position in a path finding problem.

Bridle & McCreath [7] propose a method for learning transition models in a RL agent. This reduced the number of trials required by the agent in finding an optimal policy. This was done by taking the contexts into consideration.

Flow of Events

The flow of events for the refinement process is as follows:

I The default context is activated and controls the agent.

II The Reinforced values *(Rx)* for each context, state and action triple are all initialized to zero.

III An action in the active context is carried out and this leads the agent to a new state.

IV The value of that action in that state in the current context is calculated and the *Rx* values updated based on the rewards for the mission goal -

$$Rx(c_i, s_i, a_i) = \quad (c_i, s_i, a_i) + g.\max[\, c_i s_j a]$$

V The sentinel rules are checked to see if the current active context needs to be deactivated.

   a. If a new context is needed, the listed compatible next context is activated and control returns to III.

   b. If the characteristics of the listed compatible context do not match the current situation, the *context selector* module is activated. The job of the context selector module is to search through all defined contexts to see if any matches the current situation.

      i. If there is a match, this context is activated and a copy of the previously active context is made in the *context repository*. This copy is refined by calling the *context modifier* module; the context modifier does this by adding the active context amongst the list of compatible contexts.

      ii. Control is returned to III.

   c. If none of the predefined context matches the current situation the *context creator* module is called. The context creator module would create a new context based on a predefined context template by adding the various parameters of the current situation to this template as obtained from the global fact base. Control is then returned to III.

VI If the mission goal is achieved, this marks the end of an episode, a new episode is started until the change in *Rx* values are negligible, i.e. the values converge.

VII Based on the *Rx* values, choose the action that produces the most reward for each state in a given context by choosing the max *Rx* value for each context-state-action triple combination; compare the predefined actions in a state in a context with the actions calculated for the same state in the same context.

   a. If the calculated action is different from the predefined action, create a copy of the context in the context repository and call the context modifier to refine the context with the calculated action for that state in the context.

   b. If the predefined action is the same as the calculated action, do nothing

The flow of events detailed above can refine actions within a context, context transition actions as well as create a new context based on a predefined context template.

Conclusion and Summary

This paper describes an approach to refining models of human behavior in tactical situations. It employs reinforcement learning and a simulation of the environment and missions to be undertaken by an agent controlled by the model. A prototype was built and the results indicate success.

There is ongoing work on analyzing the behavior of the agent when it is presented with a complex road network.

References

1. Oxford Dictionary; www.dictionary.com
2. Turner, E. H., Turner, R. M. “A Schema-based Approach to Cooperative Behaviour”. In Proceedings of the Thirteenth Annual Conference of the Cognitive Science Society, 1991.
3. Gonzalez, A. J., Ahlers, R. “A Novel Paradigm for Representing Tactical Knowledge in Intelligent Simulated Opponents”. In Proceedings of the 7th international conference on Industrial and engineering applications of artificial intelligence and expert systems, pp. 515-523, Austin, Texas. 1994.
4. Gonzalez, F. G., Grejs, P. & Gonzalez, A. J. “Autonomous Automobile Behaviour through Context-Based Reasoning”. In Proceedings of the 12th International Florida Artificial Intelligence Research Society Conference, pp. 2-6, Orlando, Fl. May 22, 2000.
5. Viet C. Trinh, Brian S. Stensrud, Avelino J. Gonzalez, “Implementation of a Prototype Context-Based Reasoning Model onto a Physical Platform” SAWMAS 2004
6. Wan, A.D.M., Braspenning P.J. “Context and Generalization in a Multiple Goal State Reinforcement Learning Task”. In K. van Marcke & W. Daelemans (eds.), Proceedings of the Ninth Dutch Conference on Artificial Intelligence (NAIC '97), pp. 293-302. 1997.
7. Bridle, R., McCreath, E., “Improving the Learning Rate by Inducing a Transition model” In proceeding of the Third International Joint Conference on Autonomous Agents and Multiagent Systems(AAMAS2004), Columbia University New York City, July 2004