Contextual Ranking of Behaviors for Large-scale Multiagent Simulations

(Extended Abstract)

Nidhi Parikh, Madhav V. Marathe, Samarth Swarup
Network Dynamics and Simulation Science Lab,
Biocomplexity Institute of Virginia
Virginia Tech,
Blacksburg, VA, USA.
{nidhip,mmarathe,swarup}@bi.vt.edu

ABSTRACT

As large-scale, complex multiagent simulations are becoming common, there is a need for new methods to analyze results of these simulations. One of the goals in such cases is to understand the effects of various behaviors on outcomes of interest. Here, we present a method for contextual ranking of behaviors where a partial context may already be provided in the query. Our approach uses causally-relevant states (states that have a measurable effect on the outcomes of interest), which provide the context for ranking behaviors. Apart from the partial context that may be provided in the query, our method also discovers any additional context that may affect behavioral ranking. We apply it to a large-scale disaster simulation and present results.

Keywords

behavior ranking; causal states; multiagent simulations

1. INTRODUCTION

Outcomes of disasters are driven by human behavioral responses [1]. People engage in behaviors such as looking for family members, following the crowd while attempting to evacuate, seeking information, and more, in the process of responding to any physical event. In order to develop effective plans, it is important to understand when and where people's behaviors are helpful or harmful, with the goal of channeling their natural instincts in beneficial directions.

In general, therefore, we would like to be able to *rank* behaviors in terms of their effects on outcomes. Here, we study the problem of ranking behaviors based on their effects on the final outcomes of interest for multiagent simulations. However, it may be the case that behaviors have different effects in different contexts, e.g., in the aftermath of a nuclear blast, seeking healthcare may improve an injured agent's health but if the same agent is close to the blast area and seeking healthcare early on, then it may be exposed to more radiation and this behavior can actually reduce health. Any model that does not take into account contextual information would have to take average across all contexts. This may lead to inaccurate estimation of effects and ranking. Therefore, we term this problem *contextual behavior ranking.*

By context, we mean any information (current or previous agent states and behaviors) that may lead to different outcomes for the same behavior. In our previous work [3], we proposed an algorithm to summarize simulation results by extracting causally-relevant states – states that have a measurable effect of the final outcome. These causally-relevant states are the context information mentioned above.

2. OUR APPROACH

Our goal is to contextually rank behaviors given a query which may include partial context. In addition to the partial context provided in the query, it should also be able to discover any context that may affect behavioral ranking. We use the causally-relevant states, extracted by our summarization algorithm [3], as they provide the required context. Next, we briefly describe this summarization algorithm.

It assumes that each agent trajectory is generated by the same stochastic process. Let N be the number of agents in the simulation. The state of an agent a at time t is defined by a k-dimensional state vector $\mathbf{x_a}(t) = [x_1(t), x_2(t), \ldots x_k(t)]$. Let d_i be the number of possible values that x_i can take so $d = \prod_{i=1}^{k} d_i$ is the total number of agent states. The simulation proceeds in discrete time steps from $t = 0 \ldots T$. Let the outcome variable for agent a be denoted by y_a.

In this algorithm, at each time step t, agents are divided into a set of clusters, $C(t) = \{C_1(t) \cup C_2(t) \cup \ldots C_m(t)\}$. Initially, all agents belong to one cluster. At the next time step, the state of an agent can change in d ways and so an arbitrary cluster $C_i(t)$ can split into up to d groups at the next time step $t+1$. But not all of these changes may have a significant effect on the outcome. We call each cluster derived from cluster $C_i(t)$ as candidate cluster $CC_{i,j}(t+1)$ where $j = 1 \ldots d$. To check if the candidate cluster $CC_{i,j}(t+1)$ has a significant effect on the outcome, we perform Kolmogorov-Smirnov test. Our null hypothesis is

$$Pr(Y|CC_{i,j}(t+1)) = Pr(Y|C_i(t)) \quad (1)$$

There is also a threshold δ on the "effect size", measured as a KL-divergence between $Pr(Y|C_i(t))$ and $Pr(Y|CC_{i,j}(t+1))$. If the null hypothesis is rejected and $D_{KL}(Pr(Y|C_i(t))|| Pr(Y|CC_{i,j}(t+1))) > \delta$, then candidate cluster $CC_{i,j}(t+1)$ is accepted as a new cluster at time step $t+1$. Thus, the algorithm extracts causally-relevant states – states that have a significant effect on the final outcome.

Appears in: *Proc. of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2017), S. Das, E. Durfee, K. Larson, M. Winikoff (eds.), May 8–12, 2017, São Paulo, Brazil.*

However, it may be the case that two or more individual states at consecutive time steps are not causally-relevant but the consecutive sequence of them is. For example, in a nuclear blast scenario, sheltering at any given time step may not have any significant effect, but sheltering for a long period of time may shield from radiation and lead to a significant improvement in health. Such causally-relevant sequences of states also provide the required context and hence, we adapt the summarization algorithm to also extract causally-relevant sequences of states as described below.

Our goal is to also extract causally-relevant state sequences, so while creating candidate clusters, rather than only looking at the state at the current time step (as in [3]), we look at increasingly longer past state sequences. Whenever a cluster splits from its parent cluster, any effects that the state variables upto that time step had on the final outcome has already been captured. So while creating candidate clusters for cluster $C_i(t)$, we look back only upto the time step when it was last split from its parent cluster. To check if the sequence associated with a candidate cluster $CC_{i,j}(t+1)$ affects the final outcome, we compare $Pr(CC_{i,j}(t+1))$ with $Pr(Y|C'_i(t+1-l)$ using Kolmogorov-Smirnov test, where $t+1-l$ is the length of the state sequence associated with $CC_{i,j}(t+1)$ and $C'_i(t+1-l)$ is the ancestor of $CC_{i,j}(t+1)$ at time step $t+1-l$. The "effect size" threshold δ is measured as a total variation distance between $Pr(CC_{i,j}(t+1))$ and $Pr(Y|C'_i(t+1-l)$. But some of these clusters may overlap with each other. So among the overlapping candidate clusters, we select the one that has the most effect on the outcome (i.e. the one with the highest value of total variation distance from its ancestor). If two candidate clusters have the same effect, we choose the one that requires the least amount of information about the states (i.e., the one associated with the smallest state sequence).

These causally-relevant sequences are organized in a tree structure (called causal tree), which is then matched against the query for contextual ranking of behaviors and discovering the additional context required.

The ranking algorithm takes causal tree T, a context map M, and a query q as input. Each cluster (or node) in the causal tree contains information about the state sequence associated with it, the time stamp when it was split from the parent cluster, and its score (the expected change in the final outcome that it leads to). Query q consists of time step t when we want to rank behaviors and optionally partial context (values for some of the state variables). The context map M is a hash map that maps context (time stamp and state variables) to the list of clusters matching that context. Let $qCls$ denote the set of clusters matching query q.

Our problem is to contextually rank clusters in $qCls$. The complete context for a cluster $c \in qCls$ is path from the root of tree T to c. However, some of the clusters in $qCls$ may have common ancestors and context associated with these common ancestors do not really differentiate them. Hence, to get the context that differentiate these clusters, one only needs to look back upto their least common ancestor lca.

The clusters which are siblings share the same context. So they are ranked by their scores in descending order. Let length of the context for cluster c be defined as sum over length of state sequence associated with each node n in the path from lca to its parent. So higher length gives more detailed context information. For clusters which are not siblings, they may have different lengths of contexts and it may happen that the context for one is a subset of the other. Ideally we would want clusters with more detailed context to appear first. So non-sibling clusters are sorted in descending order by the lengths of their contexts.

3. EXPERIMENTS

We apply our algorithm to a large-scale simulation of a hypothetical detonation of a nuclear device in Washington DC [4]. The simulation is comprised of a detailed, high fidelity "synthetic information system" [2] which represents the human population of the region and detailed models of four infrastructures: cell phone communication system, power system, transportation system, and healthcare system.

Agents are defined by a number of state variables. However, for the ranking purpose, we focus on six variables: agent health (modeled on a 0 to 7 range where 0 represents dead and 7 represents full health), agent behavior (six behaviors mentioned in the next paragraph, plus categories indicating if agent is in healthcare location or out of the area), if the agent has received an emergency broadcast (EBR), the agent's exposure to radiation, the agent's distance from the blast area, and if the agent has received treatment.

Agent behavior is conceptually based on the formalism of decentralized semi-Markov decision process (Dec-SMDP) with communication using the options' framework. High level behaviors are modeled as a collection of options. We model six high level behaviors: household reconstitution (HRO), evacuation, shelter-seeking, healthcare-seeking, worry, and aid & assist. These options are policies over low level actions. The actions are: to call, to text or to move. These actions are supported by infrastructural systems. Details of agent design and behavior can be found in [4].

3.1 Results

We use $\delta = 0.5$ for extracting causally-relevant states and to create a causal tree, which is used for contextual behavioral ranking. We present results from query: rank behaviors after 1 hour for people who are moderately injured.

Results show that for people who started within 0.6 mile of the blast area and with high radiation exposure, if they have managed to get further than 1 mile from ground zero, then seeking-healthcare is the best behavior, otherwise worry is the best behavior. In worry behavior, people run outside looking for information, call 911 or go to the nearest healthcare location. If the 911 call is successful, some of them may be transported to the nearest healthcare location. So worry behavior also includes some of the benefits of healthcare-seeking behavior. Since the radiation exposure is already high, reaching a healthcare location and receiving treatment is the best thing for these people to do.

In contrast, for people who started further than 0.6 miles from the ground zero and have low radiation exposure after 1 hour, all behaviors are better than worrying. Thus, our algorithm discovers different situations when the same behavior (e.g., worry) might be the best or the worst behavior.

Acknowledgments

This work has been supported in part by DTRA CNIMS Contract HDTRA1-11-D-0016-0001, DTRA Grant HDTRA1-11-1-0016, NIH MIDAS Grant 5U01GM070694-11, NIH Grant 1R01GM109718, NSF NetSE Grant CNS-1011769, and NSF SDCI Grant OCI-1032677.

REFERENCES

[1] F. Atun. Understanding effects of complexity in cities during disasters. In C. Walloth, J. M. Gurr, and J. A. Schmidt, editors, *Understanding Complex Urban Systems: Multidisciplinary Approaches to Modeling*, pages 51–66. Springer, Switzerland, 2014.

[2] M. Marathe, H. Mortveit, N. Parikh, and S. Swarup. Prescriptive analytics using synthetic information. In W. H. Hsu, editor, *Emerging Trends in Predictive Analytics: Risk Management and Decision Making*. IGI Global, 2014.

[3] N. Parikh, M. Marathe, and S. Swarup. Summarizing simulation results using causally-relevant states. In N. Osman and C. Sierra, editors, *Autonomous Agents and Multiagent Systems: AAMAS 2016 Workshops, Visionary Papers, Singapore, Singapore, May 9-10, 2016, Revised Selected Papers*, pages 88–103. Springer International Publishing, 2016.

[4] N. Parikh, S. Swarup, P. E. Stretz, C. M. Rivers, B. L. Lewis, M. V. Marathe, S. G. Eubank, C. L. Barrett, K. Lum, and Y. Chungbaek. Modeling human behavior in the aftermath of a hypothetical improvised nuclear detonation. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, Saint Paul, MN, USA, May 2013.