Solving Belief-Driven Pathfinding using Monte-Carlo Tree Search

Davide Aversa
Department of Computer, Control, and Management Engineering
Sapienza University of Rome
aversa@dis.uniroma1.it

Stavros Vassos
Department of Computer, Control, and Management Engineering
Sapienza University of Rome
vassos@dis.uniroma.it

ABSTRACT

In this work we discuss a stochastic extension to the (discrete) *Belief-Driven Pathfinding (BDP)* approach for finding personalized paths based on the beliefs of a character about the current state of the map. Our stochastic BDP upgrades previous work to the more realistic setting of using probabilities for the beliefs and takes advantage of approximate Monte Carlo Tree Search approaches.

Keywords

Pathfinding, Artificial Intelligence, Beliefs, Monte Carlo TS

INTRODUCTION

As the demand for a higher level of believability for non-player characters in games (NPCs) increases, the navigational behavior of characters should take into account what they have observed and believe about the current state of the map, instead of the actual state of affairs. This is typically sidestepped in game design or dealt in a per-case manner at the higher-level decision procedures of the NPC AI, instead of providing a general solution. A recent approach for *Belief-Driven Pathfinding (BDP)* looks into a more general game-independent methodology (Aversa and Vassos 2014). This is perhaps the only general approach for this type of pathfinding in games, and works based on a discrete "open/closed" information about some special locations on the map such as doors. In this work we extend this to the more realistic setting of using probabilities. Moreover, we connect BDP to traditional AI work in stochastic deliberation, which opens up a channel for using sophisticated AI approaches for handling uncertainty in pathfinding in games.

PATHFINDING WITH BELIEFS AND PROBABILITIES

Consider a character who has to navigate in a known map composed by several rooms connected with doors. The character knows where the doors are but does not know if they are open or closed. The doors can change state often during the lifespan of the agent. The goal of the agent is to be able to navigate the environment several times minimizing the traveling cost. In this setting, while searching for the optimal path we cannot rely solely on the path length as for instance a shorter path with low probability of having a needed door open may force the character to follow a long detour.

Proceedings of 1st International Joint Conference of DiGRA and FDG

A MONTE CARLO APPROACH

Our work with *Stochastic BDP (SBDP)* extends the Belief-Driven Pathfinding (BDP) approach (Aversa and Vassos 2014) to the stochastic setting. Initialy the map is abstracted in order to identify special locations that can be blocked/unblocked, which we call *entrances*. However, instead of storing a true/false information for the state of an entrance e, SBDP relies on a probability metric, e.g., initially $p(e) = 0.5$. This process essentially generates a *random graph* $\mathcal{G}$. Because of the stochastic nature of the graph, it is not possible to find a fixed path solution. Instead, the character needs to compute a *strategy* or *policy* that has information about the different outcomes he may see wrt the state of entrances. This problem is a special case of the *Canadian Traveler Problem (CPT)* or *Stochastic Shortest-Path Problem with Recourse* (Polychronopoulos and Tsitsiklis 1996) and can be reduced to a *Markov's Decision Process*. Following general AI solutions for CPT, we developed three approximate *Monte-Carlo Tree Search* solutions based on the *Optimistic Policy*, *Hindsight Optimization* and the *Optimistic Rollout* (Eyerich et al. 2010).

Similarly to BDP, SBDP is based on the interleaving of a *planning phase* and an *execution phase*. Once a policy is found, the character can start its execution phase in which he starts applying the policy to navigate while updating his beliefs. In particular belief update consists of two main actions: i) *sensing*: when the character can perceive the actual state of an entrance e, he updates the value of $p(e)$ to 1, if open, or 0, if closed, and ii) *knowledge decay*: this is introduced to model the natural increase of uncertainty about the state of an entrance e over time as follows: $p(t+1) = p(t) + \alpha\left[\frac{1}{2} - p(t)\right]$, where $\alpha \in [0, 1] \subset \mathbf{R}$ is the decay speed parameter that specifies how fast the probability goes to 0.5.

CURRENT AND FUTURE WORK ON EVALUATION

We are currently investigating the optimal parameters with which SBDP runs with a performance similar to the original BDP approach while providing richer behavior based on the computed policies. Moreover, we are interested in exploring two different evaluations: first, we want to ensure that SBDP will be easy to use for game developers (for instance, by integration with existing pathfinding engines) and, second, that SBPD is able to offer gameplay depth that increases player engagement. A possible positive outcome that is supported by our current evaluation is that the minimal and easy to implement MCTS solutions such as the Optimistic Policy or Hindsight Optimization can give a measurable increase in player satisfaction and agent believability.

BIBLIOGRAPHY

Aversa, Davide, and Stavros Vassos. 2014. "Belief-Driven Pathfinding Through Personalized Map Abstraction." In *Tenth Artificial Intelligence and Interactive Digital Entertainment Conference.*

Eyerich, Patrick, Thomas Keller, and Malte Helmert. 2010. "High-Quality Policies for the Canadian Traveler's Problem." In *Third Annual Symposium on Combinatorial Search.*

Polychronopoulos, George H, and John N Tsitsiklis. 1996. "Stochastic shortest path problems with recourse." *Networks* 27 (2): 133–143.