Intuitive Pathfinding of Autonomous Agents

Carsten Hahn

Mobile and Distributed Systems Group, LMU Munich
carsten.hahn@ifi.lmu.de

Abstract. This paper treats the pathfinding of autonomous agents from a given start point to a set of targets. It mainly focuses on continuous domains with continuous state and action spaces of the agents. In these continuous domains a technique called Cross Entropy Online Planning can be deployed by the agents to plan and assess sequences of actions in order to reach their goals. However, this technique requires a simulation model of the environment to enable the agents to predict and evaluate future states of the environment. This could be combined with results of the field of Intuitive Physics which tries to create models for simple physical object interactions based merely on observations. The contribution of this work is the idea to replace the simulation model in Online Planning with results of the field of Intuitive Physics in order to create Intuitive Acting Autonomous Agents. The benefit of this will be the automated creation of the otherwise manually created model and an enhancement of the flexibility of the agents.

Keywords: Artificial Neural Networks, Intuitive Physics, Online Planning, Cross Entropy, Autonomous Agents, Machine Learning

1 Introduction

Pathfinding denotes the problem of finding a path between a given start point and one or more goal points. Generally this is done with respect to a given metric in order to find for example the shortest path between two points which is quite relevant to location based services [1]. A common form of it could be a user, who shall be guided to a desired target give the knowledge of the map and his position determined by the GPS module of his smartphone. In this context the map is usually represented in a discrete manner as weighted graph and the pathfinding can be performed by using the well-known shortest path algorithms in graph theory like Dijkstra's algorithm [2].

Nowadays the instance in the need of finding a path between two points does not necessarily have to be a human. It could also be some sort of autonomous acting agent. In the recent years there is an increased interest in deploying robots in home automation or in Industry 4.0 production scenarios. It is intended that those robots act mostly autonomously which is why they also face the problem of finding paths between their current position and certain goal points. These robots can be considered as autonomous agents as they will be called in the following of this paper. They usually act on two dimensional floor plans which enables them to move in a continuous domain, except for static borders and obstacles that have to be avoided. In these continuous moving domains the previous mentioned graph theory algorithms for finding the shortest path cannot be directly applied. That is why multiple approaches exist to discretize such spaces [3] in order to be able to apply shortest path algorithms on the resulting graph. Additionally to the continuous space the robots can move in, they also face a continuous action domain which might for example consist of turning a certain degree and then moving in that direction with a certain velocity for a defined amount of time. Moreover it might also be required for the robots to adapt to changes in the environment they move in. Their desired targets might move or there might even be dynamic obstacles additional to the static obstacles that have to be avoided.

There are approaches like Cross Entropy Open-Loop Planning (CEOLP) [4] respectively Time-Adaptive Cross Entropy Planning (TACE) [5] which try to cope such complex problem domains. In TACE the agent is equipped with a simulation of the environment he is acting in. This simulation enables the agent to estimate the dynamics of its surroundings in order to plan and evaluate sequences of actions to reach its targets. This has the drawback that the environment the agent is acting in, has to be previously (manually) modeled and this model hast to be adapted if this environment changes.

On the contrary, getting a feeling for the dynamics of their environment is something humans intuitively learn while growing up. This is subject of a research field called Intuitive or Naive Physics which tries to build models of the physical interaction between objects only based on experiences respectively based on observations.

The focus of this paper is on the question to what extent the previously manually build simulation model in Cross Entropy Planning can be replaced with results from the research field of Intuitive Physics in order to enable the agent to predict the dynamics of its environment.

The contribution of this work is the idea of combining Cross Entropy Planning of autonomous agents in continuous state and action spaces with results of Intuitive Physics for automatic model building together with a discussion of possibilities, challenges and limitations.

This paper is structures as follows: In Section 2 existing approaches for pathfinding in discrete and continuous domains are outlined followed by an overview of the research filed of Intuitive Physics and how autonomous agents might benefit from results in this field in Section 3. In Section 4 the new approach of combining Online Planning with Intuitive Physics is discussed followed by a conclusion in Section 5.

2 Pathfinding of Autonomous Agents

There are numerous occasions in which an autonomous agent has to find or plan a path from a start point to a predefined goal. The strategy that is deployed by the agent in order to plan its path is bound to the representation of its environment that it has on hand. This world representation can be distinguished in discrete and continuous and will be discussed in the following section.

2.1 Discrete Domains

Discrete domains, like grid worlds, can be interpreted as graph and in the case of uniform edge weights already Depth-first or Breadth-first search are suitable to find the shortest path from a given node to a target, as long as cycles are handled accordingly. In the case of diverse edge costs Dijkstra's algorithm [2] can be used to find the shortest path from a given start node to all other nodes. Is there for example a geographical relationship between the nodes a heuristic of the shortest direct path between the start and goal can be uses to facilitate the path search, resulting in the A* search algorithm [6]. Even further, the D* algorithm [7] was developed to reduce the search costs in domains in which the graph might change dynamically. In continuous domains, on the other hand, graph search algorithms are not directly applicable. Hence multiple approaches exist for the quantization of these domains.

2.2 Continuous Domains

Considering the continuous domain as depicted in **Fig. 1**. The agent is the circle at the bottom left. It has to collect the red boxes (which might be in motion) for which it receives a reward. The walkable area is framed by grey borders and the grey rectangles are static obstacles. There might also be dynamic obstacles which follow certain movement patterns depicted as the grey circle. The action space of the agent consists of turning a certain angle and then moving in that direction for a certain amount of time.

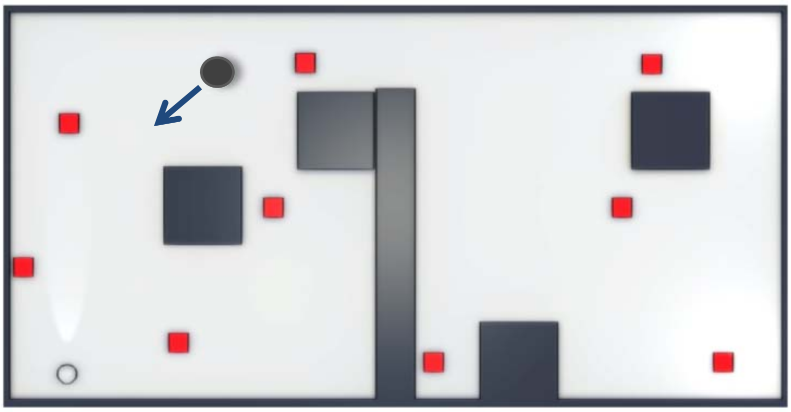

Fig. 1. Example of a continuous domain. Borders and static obstacles are depicted in grey. The grey circle represents a dynamic obstacle. The red squares represent the targets (possibly in motion) of the agent (circle bottom left). Figure taken from [8]

In previous works problem domains of this form were countered with Cross Entropy Open-Loop Planning (CEOLP) [4] respectively Time-Adaptive Cross Entropy Planning (TACE) [5]. In these approaches the agent is equipped with a simulation of the environment it is acting in. This enables the agent to simulate its actions before executing them and to aggregate quantitative information about the outcome of a sequence of actions (plan). By aggregating information about multiple plans a probability distribution over plans can be constructed where potentially promising plans are more likely to be drawn. By interleaving the execution of actions and assessment of plans by simulation the probability to draw high value plans is increased, as lined out in **Fig. 2**.

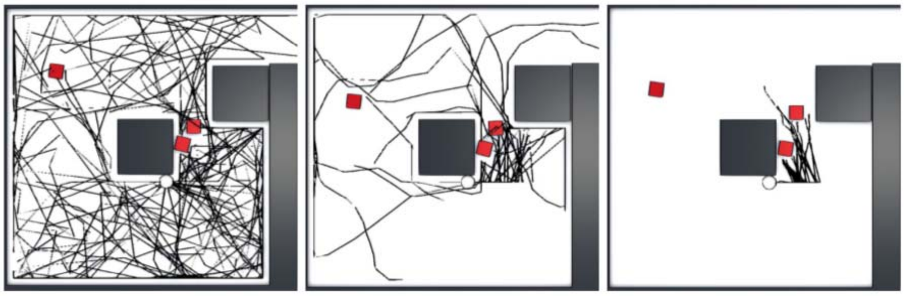

Fig. 2. Illustration of iterative variance reduction and depth adaptation of plan distributions with TACE planning. Black lines show simulation traces. While planning, the agent (white circle) identified a local problem (red boxes to be collected) and concentrated its evaluation effort accordingly. Taken from [5]

Typically, the simulation of candidate plans is the most expensive part of this procedure, which TACE tries to reduce by adaptively optimizing plan length and individual action duration. Nevertheless, the agent has to be manually equipped with a simulation of its environment. The drawbacks of this procedure are as follows:

- The simulation model of the environment has to be manually built.
- The evaluation of the simulation model could be computation-intensive especially for the simulation of physical processes.
- The pre-built model is inflexible to changes in the environment.

This is where results of the field of Intuitive Physics could put into play which will be explained in the course next section.

3 Intuitive Acting Autonomous Agents

Autonomous agents that use some sort of Cross Entropy Planning to plan and evaluate their actions in order to reach their target highly depend on the simulation model of their environment. The more accurate this model describes the dynamics, movement patterns and physical interactions of object in the environment, like moving targets, dynamic objects and their collision, the more accurate the agent can predict these processes and adapt its plan accordingly. However it can be complex and costly to develop models which accurately describe the environment the agent acts in and furthermore it might be very computation-intensive for the agent to simulate its environment using these models. Moreover these pre-built models might be inflexible to changes in the environment and could require manual adjustments.

It would be desirable that the agents had a basic understanding of object dynamics and their physical interaction just like the one that humans develop while growing up. If the agent had for example a basic understanding of concepts like bouncing objects, it could estimate the outcome of a collision without the need of exact physical simulation model and its calculation. This kind of thoughts are closely related to the field of Intuitive Physics which tries to enable artificial intelligences to predict physical relationships between objects only based on previous observations.

The field of Intuitive or Naive Physics is driven by the concept that humans have the ability to predict the outcome of physical interactions between object on a macroscopic scale only based on their experience respectively based on observations they made before. Only by observation, humans develop an understanding of concepts like gravity, object solidity, conservation of shape and momentum of the objects involved without the explicit knowledge of for example Newtonian laws [9] [10]. It is argued that an intuitive physics understanding of artificial intelligence only based on observations could be a building block for better common reasoning thereof. In the course of Intuitive Physics research multiple approaches where suggested that use artificial neural networks in order to build a model for the physical relationship that should be predicted.

Artificial neural networks can be seen as biologically inspired generic function approximators. They consist of artificial neurons that are arranged in a network of multiple layers through which an input is propagated in order to map it to a certain output. Every edge in this network is associated with a weight. These weights determine to which extent the input received over this edge influences the activation of the neuron as its activation is commonly calculated as the weighted sum of its inputs. An output function applied on the aggregated inputs determines the outcome which serves as input of the following neuron in the network or as overall output of the network in the last layer. Neural networks can be "trained" by example. This means that the error between the output of the network for a certain input and the known desired output can be minimized by propagating the error backwards through the network in order to adjust the edge weights accordingly with the goal to also minimize the output error on unseen but similar input data.

One work in the field of Intuitive Physics is [11] where the authors present a learning-based approach based on simulated data that predicts stability of towers comprised of wooden blocks only based on an image as input. The model was trained with simulated data but tested with a real robot, which placed wooden blocks as a tower and was guided by the model for the prediction of the stability of future block tower states.

In another work the concerning authors considered the domain of playing billiards [10]. The authors suggest an object based, image centered model to predict the movement and collision of billiard balls. The input for the model is a stack of 4 images comprised of the current as well as the previous 3 time steps focused on the object (billiard ball) to be predicted as well as the force that is applied to it at the current time step. The model then predicts the velocity of the object at each of h future time steps, with $h = 20$ for example. This prediction of future states is done for each object individually to predict the future world state, while the same model is applied to all the objects in the world. To explore the generalization of this method, the predictive model was trained on a variety of billiards environments, with different numbers of balls and different wall geometries and later tested on unseen setting. The authors argue that it could be valuable for an intelligent agent to be equipped with an internal model of the dynamics of the external world in order to plan and execute goal specific actions in varied and unseen environments.

In order to further investigate the possibility of predicting the future state of dynamic objects with a model that is only build on previous observations, an additional experiment of Intuitive Physics was conducted in an unpublished preparatory work of the author of this paper.

The scenario consisted of an elastic ball that bounced in an enclosed two dimensional area after being initialized at a random position with an impulse in a random direction. The ball was subject to gravity and friction which slowed it down. The goal in this scenario was to build a model to predict the trajectory of the ball after a certain point after being feed with a segment of the previous part of the path. The model was created only by observing trajectories of the ball. For this reason 100 000 example trajectories where created by using a physics simulation. After that the dataset was split into training and test set. The input of the model is the x and y coordinate in the area as well as the velocity in x and y direction for the last 50 time steps, which means that the used artificial neural network had 200 input neurons. The neural network was used to predict the next 20 time steps of the ball trajectory. The architecture of the neural network was a simple feed forward network with 4 fully connected layers each consisting of 200 neurons with rectified linear unit (ReLU) activation. These fully connected layers where interleaved with dropout layers (20%) for regularization. As loss function the mean squared error between the models output and the real measurements from the physics simulation was used for training.

The evaluation on the test set showed promising results concerning the accuracy of the prediction of the ball trajectory. This preliminary investigation as well as the other related work of Intuitive Physics lead to the assumption that an autonomous agent could benefit from the possibility to predict the dynamics in its surround in order to find, respectively plan, its path to (possibly moving targets) while avoiding static and dynamic obstacles. The idea is to replace the simulation model in the previous mentioned Cross Entropy Planning with (possibly pre-trained) neural networks to enable the agent to estimate the dynamics of its surrounding. The possibilities and challenges of this approach shall be discussed in the following section.

4 Discussion

In previous sections the technique of Cross Entropy Planning was explained. In this approach an autonomous acting agent is equipped with a simulation of its environment in order to plan and assess sequences of actions. These actions could consist of turning a certain angle and moving forward in the facing direction with a certain speed for a defined amount of time. Action sequences that lead the agent to a predefined goal yield higher expected values and are therefore chosen for execution. In a dynamic environment also the dynamics of this environment have to be modeled as accurate as possible. One big drawback is that this simulation model has to be manually built which might be labor-intensive. It would be desirable that the agent had an intuitive concept about the dynamics in its environment and could estimate respectively predict the dynamics of its targets or dynamic obstacles. This is where results of the field of Intuitive Physics could put into play, as researchers of this field try to automatically build models for physical processes merely bases on observations. This lead to the idea that the agent could be provided with artificial neural networks to predict the dynamics of its surroundings, may it be targets or obstacles. This approach shall ease the creation of the model and make it more flexible to changes in the environment. Moreover the evaluation of the neural network could be less computation-intensive than the evaluation of a complex physics simulation, depending on the complexity of the model and the environment. The neural networks could be pre-trained in other domains in order to model physical object relations. It needs to be examined whether these models can be further trained in the target domain in order to adapt to changes or whether the training of the neural networks for dynamic object prediction could start together with the plan generation of the agent respectively be interleaved with the plan generation. It is to be further investigated how well neural networks as future state estimator perform compared to detailed physical simulations in terms of efficiency and accuracy of the prediction and to which extent this approach can be generalized.

5 Conclusion

In this work the concept of Intuitive Acting Autonomous Agents was introduced. The idea is to replace the simulation model in Cross Entropy Online Planning with results of the field of Intuitive Physics and combine these research areas thereby. Promising results in Intuitive Physics show that it is possible to build models for simple physical object interactions only based on observations. This might ease the model creation and flexibility of the model. Further research is required to investigate how well neural networks perform compared to a manually defined simulation of the agent's surroundings in order to plan and assess future states.

References

1. Hannah Bast et al., "Route Planning in Transportation Networks," CoRR, vol. abs/1504.05140, 2015.
2. E. W. Dijkstra, "A note on two problems in connexion with graphs," Numerische Mathematik, vol. 1, no. 1, pp. 269-271, Dec 1959.
3. I. Millington and J. Funge, Artificial Intelligence for Games.: CRC Press, 2016.
4. Ari Weinstein and Michael L Littman, "Open-Loop Planning in Large-Scale Stochastic Domains.," in AAAI, 2013.
5. Lenz Belzner, "Time-adaptive Cross Entropy Planning," in Proceedings of the 31st Annual ACM Symposium on Applied Computing, New York, NY, USA, 2016, pp. 254-259.
6. P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths," IEEE Transactions on Systems Science and Cybernetics, vol. 4, no. 2, pp. 100-107, July 1968.

7. A. Stentz, "Optimal and efficient path planning for partially-known environments," in Proceedings of the 1994 IEEE International Conference on Robotics and Automation, May 1994, pp. 3310-3317 vol.4.
8. Lenz Belzner, "Simulation-based autonomous systems in discrete and continuous domains," LMU Munich, Ph.D. dissertation 2016.
9. Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman, "Building Machines That Learn and Think Like People," Behavioral and Brain Sciences, pp. 1–101, 2016.
10. Katerina Fragkiadaki, Pulkit Agrawal, Sergey Levine, and Jitendra Malik, "Learning Visual Predictive Models of Physics for Playing Billiards," CoRR, 2015.
11. Wenbin Li, Ales Leonardis, and Mario Fritz, "Visual Stability Prediction and Its Application to Manipulation," CoRR, vol. abs/1609.04861, 2016.