Physics-Engine-Based Robot Planning in Physics Projection

Kosuke Nishie Yoshiki Takatsu Naoto Iwahashi

Okayama Prefectural Univ

We have earlier proposed a scheme, named *physics projection*, for artificial intelligence systems to make sequential decisions in physical environments. In this paper, we present a novel physical planning method in physics projection for robots. It has two original features: 1) online uncertainty reduction and 2) adaptive model predictive control. In the method, an adaptively learned physical world model is used as the predictor of the action effects in the physical world. The results of the preliminary experiments using a physics simulator are shown.

1. Introduction

Physical planning is at the core of human intelligence. Physical planning has long been studied in the broad fields of science and technology [Craik 63, Polanyi 66, Winograd 72, Gibson 79, Brooks 91, Spelke 07, Battaglia 13, Lake 17, Kunze 17, Watters 17, Northoff 18, Toussaint 18, Kloss 18, Sünderhauf 18, Janner 19, Bakhtin 19]. *Physics projection* (Fig. 1) was proposed by Iwahashi [Iwahashi 19] as a way for artificial intelligence systems to make sequential decisions in physical environments, and is characterized as

- Online learning of physical world model,
- Dynamic planning by simulation,
- Embodiment.

These characteristics are the same as those of human physical planning abilities, but they have not been fully implemented in artificial intelligence systems thus far. In this paper, we present a physical planning method that can be executed in physics projection.

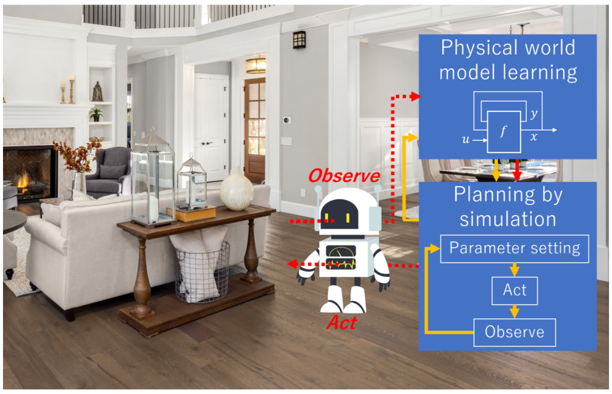

Figure 1: Physics Projection

2. Proposed Physical Planning Method

In general, physical planning is sequential decision-making, in which subsequent actions often depend on the effects of actions that precede them [Lyu 19]. We have to consider that actions can change the physical world. Considering such characteristics, we propose a novel physical planning method that has the following original features:

Online uncertainty reduction In the physical planning problem, two kinds of uncertainties have to be managed. One is uncertainty due to lack of observation, and the other is uncertainty due to unpredictability in the physical world. Both kinds of uncertainties are reduced during action control and planning by the robot's actual and simulated active observations in the physical world and physical world model, respectively.

Adaptive model predictive control The scheme of adaptive model predictive control (e.g. [Lu 19]) is adopted. The optimality of plans is estimated in the prediction horizon based on the physical world model, which is learned or adapted in an action control feedback loop to reduce its uncertainty.

Table 1: Notation used in Algorithm 1

ω	plan
u	action
x	state of the world
Ω	set of plans
$G(\omega)$	present value of ω
$H(\omega)$	expected future value of ω
$p(x\|\omega, u)$	probability that x occurs after u following ω

The algorithm of our physical planning method is presented in Algorithm. 1. Table 1 shows the notations used in the algorithm description. The physical world model provides the action-effect predictor $p(x|\omega, u)$. We can use any physical world model that is manually developed (e.g. Unity) or one that is learned by deep learning [Janner 19, Bakhtin 19].

3. Experiments

We implemented the proposed physical planning method using the physical world simulator Unity, which is composed of a graphics engine, physics engine, and fluid simulator. Although the physical planning method should be run in

Contact: iwahashi, iwahashi@c.oka-pu.ac.jp, www-ail.c.oka-pu.ac.jp

Algorithm 1 Physical planning in physics projection

Input: Physical world, an executable action set, initial physical world model, initial status x_0, goal condition GC

Output: Action sequence executed by a robot, $\hat{\omega}$, that satisfies GC

Initialisation : $\hat{\omega} = ((\phi, \hat{x}_0))$

LOOP Process

while $\hat{\omega}$ does not satisfy GC **do**

  $\omega_0 \leftarrow \hat{\omega}$, $\Omega = \{\omega_0\}$, $G(\omega_0) = 0.0$

  Prediction by the physical world model

  while ω_i does not satisfy GC for all $\omega_i \in \Omega$ **do**

    Get $p(x'_{i,j,k}|\omega_i, u_{i,j})$ for all following possible state $x'_{i,j,k}$ after action $u_{i,j}$ following plan ω_i

    $\hat{i}, \hat{j} = \text{argmax}_{i,j} \sum_k \{p(x'_{i,j,k}|\omega_i, u_{i,j}) \times (G(\omega_i + (u_{i,j}, x'_{i,j,k}) + H(\omega_i + (u_{i,j}, x'_{i,j,k}))\}$ for all $\omega_i \in \Omega$ and possible action $u_{i,j}$

    for all $x'_{\hat{i},\hat{j},k}$ **do**

      $\omega_{|\Omega|} = \omega_{\hat{i}} + (u_{\hat{j}}, x'_{\hat{i},\hat{j},k})$

      Add $\omega_{|\Omega|}$ to Ω

    end for

  end while

  Actual action execution in the physical world

  Execute the first action that is not executed, u, in ω ($\in \Omega$) satisfies GC, and get the next actual state x

  Observe the physical world

  $\hat{\omega} \leftarrow \omega + (u, x)$

  Adaptation of the physical world model

  Compare the predicted and actual effects of action u

  Renew the physical world model

end while

return $\hat{\omega}$

a real physical world, it was run under a simulated physical world build with Unity. The physical world model was also represented by Unity, and learned incrementally from the simulated physical world using *yolo* object recognition (Fig. 2).

Two simple tasks, find-and-carry and move-and-grab, were conducted using the physical planning method.

3.1 Find-and-carry

The goal condition is for the robot to find a mobile phone in a room located in the house and carry it to the user. It is assumed that the robot is aware of the probability of the mobile phone being in a particular room, and the position of the user. The execution process of the planning method is shown in Fig. 3. The robot prepared an optimal plan in the physical world model before it began to move. In the left figure, the robot moved to the room it planned to visit first, found the mobile phone, and carried it to the user in the physical world. In the right figure, the robot moved to the room it planned to visit first, but did not find the mobile phone. Then, the robot adapted the physical world model and prepared an optimal plan with the adapted physical world model. Then, the robot moved to the room it planned to visit next, where it found the mobile phone and carried it to the user in the physical world.

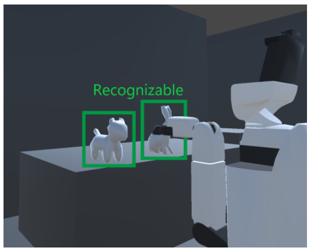

Figure 2: Representation of embodied physical world model. The objects that were recognized and not recognized in the physical world were represented by the adapted categorical objects and cubes, respectively, in the physical world model, which included a self body.

Figure 3: Physical planning for find-and-carry task.

3.2 Move-and-grab

The goal condition is for the robot to raise the toy without spilling a glass of wine placed next to the toy. The execution process of the planning method is shown in Fig. 4. The optimal plan was prepared before the robot proceeded with the activity. The top figure shows the case where the method could devise a plan that satisfied the goal condition. At time t, the robot attempted to grab a toy rabbit from the front and spilled a wine glass in the physical world model. At time $t+1$, the robot moved to the side, observed the side of the toy in the physical world, and adapted the physical world model. At time $t+2$, the robot grabbed a toy successfully in the physical world model in the physical world model. This is an optimal plan. The bottom figure shows the case where the method could not devise a plan that satisfied the goal condition. At time t, the robot attempted to grab a toy dog from the front, and spilled a wine glass in the physical world model. At time $t+1$, the robot moved to the side in the physical world, observed the side of the toy in the physical world, and adapted the physical world model. At time $t+2$, the robot attempted to grab the toy but could not grab it owing to its width, and spilled the wine glass in the physical world model. Thus, the plan that satisfies the goal condition was not devised.

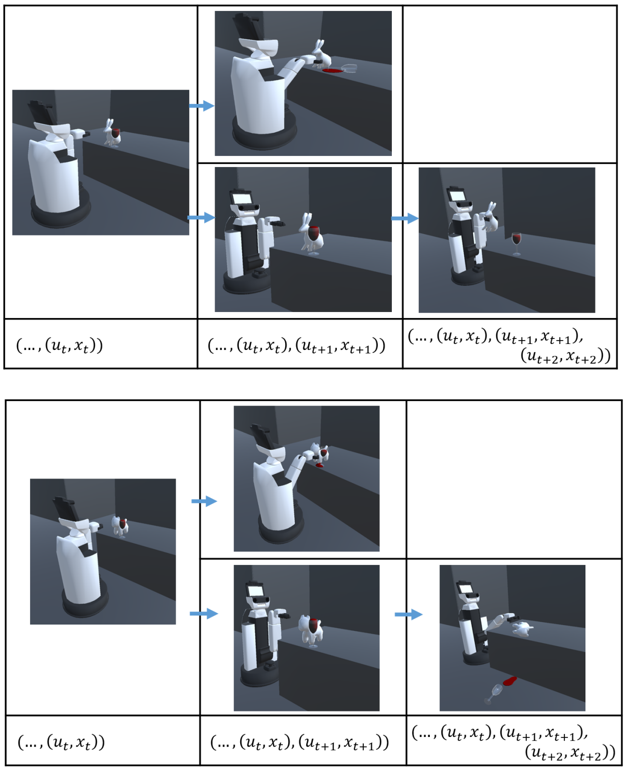

Figure 4: Physical planning for move-and-grab task. **Top:** The method could devise a plan that satisfied the goal condition. **Bottom:** The method could not devise a plan that satisfied the goal condition.

4. Conclusion

It was confirmed that the process of interaction between the physical world and the physical world model in the proposed physical planning method was successfully implemented. We need to investigate the search efficiency of the method in the future.

The proposed method is generic and can be applied to any physical task performed by robots. In addition, it can be integrated with model-based reinforcement learning to cope with environment shift problems in reinforcement learning.

References

[Bakhtin 19] Bakhtin, A., Maaten, van der L., Johnson, J., Gustafson, L., and Girshick, R.: Phyre: A new benchmark for physical reasoning, in *Advances in Neural Information Processing Systems*, pp. 5083–5094 (2019)

[Battaglia 13] Battaglia, P. W., et al.: Simulation as an engine of physical scene understanding, *Proceedings of the National Academy of Sciences*, pp. 18327–18332 (2013)

[Brooks 91] Brooks, R. A.: Intelligence without representation, *Artificial intelligence*, Vol. 47, No. 1-3, pp. 139–159 (1991)

[Craik 63] Craik, K.: *The nature of explanation*, Cambridge University Press (1963)

[Gibson 79] Gibson, J. J.: The ecological approach to visual perception. (1979)

[Iwahashi 19] Iwahashi, N.: Physics projection, in *International Conference on Awareness Science and Technology* (2019)

[Janner 19] Janner, M., Levine, S., Freeman, W. T., Tenenbaum, J. B., Finn, C., and Wu, J.: Reasoning About Physical Interactions with Object-Centric Models, in *International Conference on Learning Representations* (2019)

[Kloss 18] Kloss, A., et al.: Combining learned and analytical models for predicting action effects from sensory data, *arXiv* (2018)

[Kunze 17] Kunze, L. and Beetz, M.: Envisioning the qualitative effects of robot manipulation actions using simulation-based projections, *Artificial Intelligence*, Vol. 247, pp. 352–380 (2017)

[Lake 17] Lake, B. M., Ullman, T. D., Tenenbaum, J. B., and Gershman, S. J.: Building machines that learn and think like people, *Behavioral and Brain Sciences*, Vol. 40, (2017)

[Lu 19] Lu, X., Cannon, M., and Koksal-Rivet, D.: Robust Adaptive Model Predictive Control: Performance and Parameter Estimation, *arXiv preprint arXiv:1911.00865* (2019)

[Lyu 19] Lyu, D.: Knowledge-Based Sequential Decision-Making Under Uncertainty, in *International Conference on Logic Programming and Non-monotonic Reasoning*, pp. 3–7 (2019)

[Northoff 18] Northoff, G.: *The spontaneous brain: from the mind-body to the world-brain problem*, The MIT Press (2018)

[Polanyi 66] Polanyi, M.: The Tacit Dimension (1966)

[Spelke 07] Spelke, E. S. and Kinzler, K. D.: Core knowledge, *Developmental Science*, Vol. 10, No. 1, pp. 89–96 (2007)

[Sünderhauf 18] Sünderhauf, N., et al.: The limits and potentials of deep learning for robotics, *The International Journal of Robotics Research*, Vol. 37, No. 4-5, pp. 405–420 (2018)

[Toussaint 18] Toussaint, M., et al.: Differentiable physics and stable modes for tool-use and manipulation planning (2018)

[Watters 17] Watters, N., et al.: Visual Interaction Networks: Learning a Physics Simulator from Video, in *Advances in Neural Information Processing Systems 30*, pp. 4539–4547 (2017)

[Winograd 72] Winograd, T.: Understanding natural language, *Cognitive psychology*, Vol. 3, No. 1, pp. 1–191 (1972)