International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received**: 12 January 2023 **Revised**: 20 February 2023 **Accepted**: 27 February 2023*

An Investigative Analysis using 'Prior' Knowledge of Physical Interaction to Learn State Representation in Robotics

M. Mohammed Thaha[1], M. Preetha[2], K. Sivakumar[3], Rajendrakumar Ramadass[4]
[1]Assistant Professor (Sr.Grade), B.S.Abdur Rahman Crescent Institute of Science and Technology, GST Road, Vandalur, Chennai - 600 048, Tamilnadu, INDIA
Email: mohammedthaha@crescent.education
[2]Professor, Prince Shri Venkateshwara Padmavathy Engineering College,
Email: preetha.m.cse@psvpec.in, smpreetha14@gmail.com
Orcid: 0000-0001-8483-9871
[3]Professor, Department of Mechanical Engineering,
P.T. Lee Chengalvaraya Naicker College of Engineering and Technology, Kanchipuram
Mail Id: shivakees@gmail.com, ORCID: 0000-0002-9338-3519
[4]Asst.Trainer, Electrical Engineering Section
College of Engineering and Technology, University of Technology and Applied Sciences, Shinas - Oman
Email: rajendrakumar.ramadass@shct.edu.om

Abstract: The effectiveness of learning in robots is heavily influenced by state representations. In turn, physics gives structure to both the world's largest changes and the manner wherein robots may influence them. Using prior knowledge of engaging with the material realm, robots may develop state descriptions that are consistent with mechanics. Six mechanical priors were discovered, along with a description of how they can be used for language modelling. We demonstrate the effectiveness of our technique inside a virtual slots auto racing game and a virtual navigating assignment involving disturbing motion information. Our method extracts mission condition models from elevated observations even when task-irrelevant diversions are prevalent. We also show that the state representations learnt by our technique significantly increase reinforcement learning generalisation.

Keywords: State representations, physical world, prior knowledge, simulated navigation task, robotic priors, high- dimensional observations, reinforcement learning

I. Introduction

In robotics and artificial intelligence, a long-term goal is to develop versatile robotics capable of performing a variety of jobs on their own. Perception and learning, on the other hand, are challenged by these high-dimensional observations. This appears pointless, because most tasks can be mastered by focusing just on the components of the high-dimensional information that are relevant to them. To develop task-general robots, just those features relevant to tackling the job at hand must be extracted from the high-dimensional sensor input.

Feature engineering is perhaps the most popular approach to this problem in robotics. Hand-drawn maps from observations to state representations are created utilising. However, the disadvantage of this technique is that in order to achieve our initial aim, we must establish an observation-state mapping for each robotic activity.

Machine learning, rather than human intuition, is used in representation learning approaches to extract relevant information from high-dimensional data. This method does not need any prior understanding of the task. Instead, it makes broad assumptions about the problem's structure. However, the enormous quantity of data and processing necessary to derive effective state representations comes at a cost. Data collection is time-consuming and expensive in robots. As a result, conventional representation learning methodologies may be challenging to implement.

Robots, on the other hand, are not required to solve the problem with general transfer learning. In order to connect with the physical environment, robots merely need in descriptions. Physics structures the both difference in the society as well as the methods in which robots could affect these adjustments.

In this paper, we describe five robotic priors and illustrate how they may be used to learn state representations via translating it to an error function and shrinking it. We put our strategy to the test in two simulated robotic tasks relying on observation: an on only race car as well as a navigating job with a robot manipulator in a setting with shifting irrelevant features (see Figure 1). In all instances, its robot develops a distance matrix from reflectively image input to cheap situations. We illustrate that resulting state description adequately captures the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received**: 12 January 2023 **Revised**: 20 February 2023 **Accepted**: 27 February 2023*

important characteristics of the job, as just a consequence, domain adaptation efficiency.

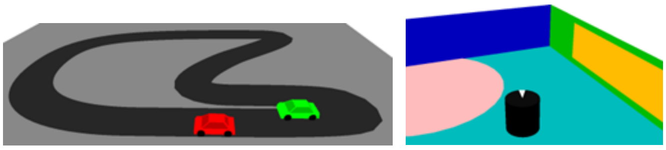


Figure 1.Visual distracters with robotic tasks by simulation

II. Literature review

The term "prior" relates to a prior confidence interval, which is then divided by the difficulty and standardized to produce the posterior. Inside the area of transfer learning, we just use word prior in a similar manner to the others [1]. We'll now look at a few other areas in which priors can be established.

The goal of representation learning research is to minimise the requirement for feature engineering by automatically selecting useful patterns from images. The performance of the proposed method has been demonstrated in problems such as speech synthesis [24], item recognition [13], and computational linguistics [4]. All of these instances vastly outperform the best existing designed representations-based approaches. To attain these outcomes, the representation learning approaches make use of generic priors, large data, and vast computing. "several broad priors about environment surrounding." They prepared a number of general AI priors, believing that by enhancing and merging this list into a representation learning technique, we can get nearer to machine learning. In the subject of robotics, this really is exactly what we are attempting. We focus on robotic challenges that include interaction with the physical environment in order to obtain stronger priors regarding the issue structure. Such priors are referred to as robotic priors.

State representation learning

It's an example of interactive representation learning, with the goal of identifying a map between observable to state which enables the proper actions to be chosen. This difficulty is more difficult than the conventional matrix factorization difficulty, that is addressed by inter scale [14] and other algorithms [23, 29, 6] since they need awareness of lengths or spatial relationships among data points within state space. The machine, on either hand, has had no previous understanding of sensory data's semantic similarity. That must solve the supervised learning problem n figuring out which observation correlate with similar events in terms of the goal, itself which would be unable to do without a decent state description. What makes a good goal, you might wonder?

Compression of Data: Lange et al. [15] use deep auto encoders to compress observations to produce state representations. The strategy is premised on the idea that the observations can be condensed into a simple (near zero) state description. While we keep things as simple as possible, we believe it's important to think about timing, activities, and consequences.

This method was used to find a description of the human machine's body postures [7] as well as to tackle reinforcement learning challenges based on visual data.

The challenge of learning a policy to pick activities in order to maximise future rewards is known as reinforcement learning [28]. The policy connects states to actions. However, because the robot can seldom observe its present state st directly, it must use an observation-state-mapping to compute st = (ot) from its observation ot (see Figure 2). The robot executes action at = st if st is given (st). This framework explains the robot's interaction with the outside environment. As a result, it's well-suited to formalising many robotics learning challenges.

State depiction learning is an effective task that involves connecting facts to states so designed to facilitate rational policy learning. This is the problem that this study attempts to solve in a robotics-specific way.

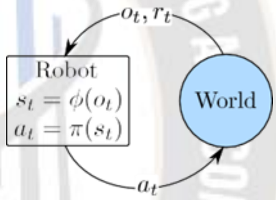


Figure.2. the robot-world interaction is depicted in Figure 2. Using observation-state-mapping, the robot computes the state st from its observation ot at time t. It takes action in accordance with policy in order to maximise future benefits rt+1

1. However, in this study, we assume that the situation is completely observable, and that the final observation has all of the information needed to determine the appropriate action. This is a significant drawback, as many real-world robotics issues may only be seen in part. Some of the constraints can be overcome by combining sensor inputs from various time steps, but not all of them.

    1. Automated Priorities

       The laws of physics govern the interaction between the robot and the actual environment. We can derive robotic priors that capture characteristics of robotic tasks from this fact.

    2. Simplicity Priority: Only a few world attributes are relevant for a particular job. This presumption is connected to Occam's razor, which is an element of the scientific approach for gaining knowledge about

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received**: 12 January 2023 **Revised**: 20 February 2023 **Accepted**: 27 February 2023*

our physical reality. It prefers state representations that exclude extraneous data, resulting in a lower-dimensional reinforcement learning problem and better generalisation.

3. Coherence in Time Prior: The world's task-relevant characteristics evolve gradually over time. Newton's first law of motion is connected to this antecedent. Physical things have inertia, and external influences only gradually modify their motion. However, because most changes in the world happen gradually, temporal coherence can apply to more abstract aspects than physical motion. As the robot transitions between states, the temporal coherence prior favours state representations that follow this approach.
4. Proportionality Prior: The magnitude of a change in task-relevant attributes as a result of an action is proportional to the action's magnitude. F = m a, Newton's second law of motion, gives us this antecedent. The acceleration produced by an action is constant if it reflects the application of a given force on a fixed mass object. This is true for mobility and physical interactions with the world's objects, but it also applies to more abstract qualities. The proportionality principle in state representation is enforced by this prior.
5. Causality Prior: The reward is determined by the task-relevant attributes along with the activity. This and the preceding laws are similar to Newton's third law of motion, or, more broadly, causal determinism. If the same action results in different rewards in two separate contexts, both scenarios must have some task-relevant attribute in common and so should not be represented by the same state. As a result, state representations that have the required features to differentiate these scenarios benefit from this prior.
6. Repeatability Prior: The task-relevant qualities and the action determine the change in these attributes as a result of the action. This prior is similar to the last one, only it is for states rather than rewards, and it is likewise based on Newton's third law of motion. This concept is maintained by choosing state representations in which actions have comparable outcomes when repeated in similar circumstances.

Even in the physical world, however, there are counter-examples to each previous example. In reality, similar counterexamples may be found in this paper's simulated robotic experiments: Reasonableness doesn't really hold when the robot crashes into such a barrier and its position remains fixed despite attempting to progress at a specified rate. The bulk of these priors are represented by a set of acts and rewards, which is worth noting. As a consequence, they are useless for "robots" that can only watch rather than act. though, because process of modeling can differ significantly to our own, for as by not following Newton's second law.

Methods

We'll now explain how minimising the loss function may be used to learn a linear mapping from observations to states.

Considerations in terms of computation: In the loss function, we compute the expected values by computing the average of training samples. For symmetry, causality, and repeatability losses, this would necessitate assessments of all O(log n) pairings of training samples. To approximate these comparisons, we only examine items that are k time step apart for computing efficiency it only has to be high enough for events separated by this degree and direction of discrepancy to be roughly mutually independent. Designers used k =1 100 in all of our studies.

In this part, we use 300-dimensional visual data to test our technique in simulated robotic tasks. To begin, we examine learnt state representations to acquire a better understanding of our approach's potential. We start comparing acquired feature models for a simple navigation challenge in which the robotic perceives the scene from egocentric or top-down viewpoints. Regardless of the fact that it is crucial, it's the first occasion that this issue is handled with state language modeling, apart in carefully controlled experiments. On various state representations, we compare a conventional reinforcement learning strategy. When compared to alternative baselines, the experiment reveals that our technique can significantly increase reinforcement learning performance.

To check if our approach is invariant to perspective, we put it to the test in multiple versions of a basic navigation challenge with separate visual observations, witnessing the terrain first from top or viewing it all from the machine's point of view. In all implementations, its robotic acquires a value description that describes its orientation, and that is exactly the learning required to perform the task.

Experiment Setup: For both versions of the task, The following experiment was carried out. The robot analysed its environment based on this experience by doing 5,000 randomized movements and establishing a translation as from deeply perception field to a this double simulation. For both state representations: the state representations have two orthogonal axes that correspond to the robot's coordinates. These state space axes, of course, do not have to line up between tests; they may be rotated or reversed. The robot's

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
***Article Received**: 12 January 2023 **Revised**: 20 February 2023 **Accepted**: 27 February 2023*

sensory observations were considerably different in the two versions of the test. It did, however, uncover a relationship among these data as well as the mission dimensions—the robot's location. In this experiment, we see if our technique can tell the difference between task-relevant features of observations and irrelevant data. In a two-vehicle slot car racing assignment, we study this.

III. Results

To compare the learned state representations, researchers displayed average forth is for these 5,000 clock cycles again for upper viewpoint and the egocentric view. The data in state vector form a square in both cases, meaning that the state provides an approximation of a machine's position in the rectangular box. We can establish that this is what is learned simply shading every condition sampling as per the machine's underlying data x-coordinate or measurement data y-coordinate.

The robot has no idea which automobile is appropriate for the task.

The Slot Car Racing Challenge is as follows: Figure 4 depicts an example scene from this challenge. Every time step, the robot may control the speed of the red car by picking [0.01, 0.02,..., 0.1] units. With such a standard error of 10% of a needed velocity, there's really minimal mean Noise present. The machine's payout is equivalent to the stated velocity except if the vehicle goes too rapidly in a steep bend and is thrown off the track. Its robot is given a financial motivation of ten dollars in this situation. The green slot vehicle is beyond the robot's control. This car's speed is picked at random from the same range as the red car's. The robot's reward or the red vehicle's mobility are unaffected by the green slot car. The robot looks down on the scene through a 10-pixel RGB picture from above.

In previous testing, we'll look at what happens whenever the bot comes to change feature representation with much more qualities than that are necessary in the same instances in this section. With egocentric observations, we reproduced the findings again for slot car job as well as the simple navigation challenge. Instead of using a two-dimensional state representation, we used a five-dimensional subspace. After probing for 5,000 number of iterations and obtaining state descriptions, we obtained a 5,000 5-matrix M holding the events.

Identifying the Problem Dimensionality: We can see this level by projection all state sample onto their first three main components. These condition values form a rectangle in this area, just as they did inside the double test. Despite having a four subspace, the robotic understands that the task is one and simply stores all features of its observations.

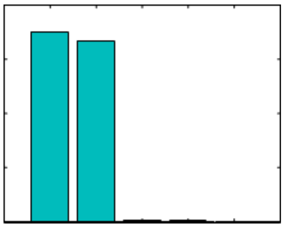


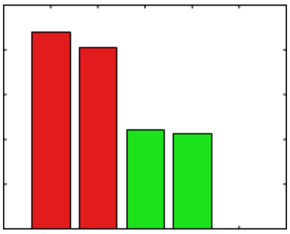
Figure 3 shows the results of a under space vector navigation problem.

Why is green car inside the state when it has nothing to do with the task? Although it performed the same task, the robot's attitude differentiation between situations grows as it receives different incentives. If indeed the robotic selects the same speed but the slot cars is pushed off one time but returns to the racetrack it's next, it tries to distinguish between the two states. The location of the red slot vehicle is the most striking distinction between these circumstances. On either hand, little differences in geography or the probabilistic reasoning of the acts might indeed make a contrast between the different results. As nothing more than a consequence, the bot seeks for other answers, such as how the blue slots car is. This characteristic's eigenvector indicate. As a consequence, the bot seeks for other answers, such as where the blue slot car is. If the state space has enough dimensions, our technique covers these alternate explanations. When the state space is constrained, the technique concentrates on the most important dimensions.

The previous trials have shown that our technique has promising qualities. However, the value of state representations can only be assessed in terms of how they aid later learning. We will show in this experiment that our technique can considerably improve reinforced academic achievement while using a small amount of data.

The task's goal has remained unchanged. The robot must get inside Fifteen meters of the upper right side to receive a result of 10, until it collides with a barrier, in which instance it will receive a score of 1.

Experimental Setup: As in prior tests, the robot explored at random but was paused every 500 time steps. It learnt an observation-state-mapping and a policy from its gathered experience. The robot was then put through its paces for 20 episodes of 50 steps in order to calculate the average reward sum. This learning-evaluation cycle was performed ten times.



The bigger eigenvalues in Figure.4 correlate to the racing track location of the controlled red slot vehicle. The location of a quasi green slot vehicle corresponds to the 3rd and 4th major components.

The raw 300-dimensional observation, this same five slightly slower characteristics of an observations (calculated utilising sequential slow feature analysis [30], the very first 5 main elements of a observations, and the four government portrayal did learn with our method were all tested between several state depictions using the same validation active learning. We also compared this task to a reduced version without misdirections, wherein the robots has knowledge to its regression coefficients posture and may establish an arbitrary cap on reinforced academic performances. It employs its location as state, as well as the cosine and sine of its orientation, which we think to be the best representation for this purpose.

These findings show that when learning algorithm is given to the variables acquired by our method, the robotic form large less training. What is the source of this discrepancy? Our approach is essentially a regularisation of the learning issue. As a result, generalisation occurs more quickly.

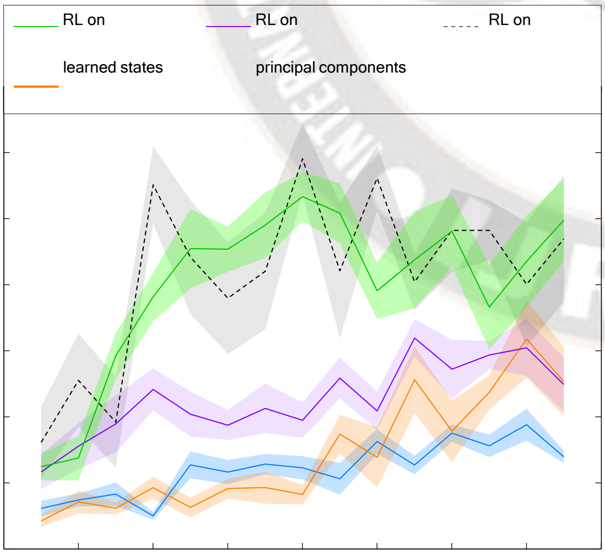


Figure 5 shows the performance of reinforcement learning for several state representations. The means are shown by lines, whereas the standard errors are represented by surfaces.

Finally, we compare our findings to the top bound of the reinforcement learning method—using the robot's ground truth stance as state (dashed line, see Figure 5). The results are comparable even after only a few hours of training, indicating that our method requires fewer data than the reinforcement of the learning.

Conclusion

We've given a method for learning state representations in robotics that is based on past knowledge of how to interact with the real environment. We may apply robotics-specific prior knowledge by reducing the issue domain in this way.

The second crucial concept is to utilise this information to assess representations based on how well they match our prior assumptions about the world. We presented five robotic priors for state representation learning: repeatability, proportionality, temporal coherence, causality, simplicity and illustrated how they may be used to achieve a goal.

In future study, we would want to recommend forming more priors on underlying issue structure in robotic priors into the knowledge pool of machine learning.

References

1. Yoshua Bengio, Aaron C. Courville, and Pascal Vincent. Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8):1798–1828, 2013.
2. Byron Boots, Sajid M. Siddiqi, and Geoffrey J. Gordon. Closing the learning-planning loop with predictive state representations. International Journal of Robotics Re- search, 30(7):954–966, 2011.
3. Michael Bowling, Ali Ghodsi, and Dana Wilkinson. Action respecting embedding. In 22nd International Conference on Machine Learning (ICML), pages 65–72, 2005.
4. Ronan Collobert, Jason Weston, Le´on Bottou, Michael Karlen, Koray Kavukcuoglu, and Pavel Kuksa. Natural language processing (almost) from scratch. Journal of Machine Learning Research, 12:2493–2537, 2011.
5. Siegmund Duell, Steffen Udluft, and Volkmar Sterzing. Solving partially observable reinforcement learning problems with recurrent neural networks. In Neural Networks: Tricks of the Trade, volume 7700 of Lecture Notes in Computer Science, pages 709–733. Springer Berlin Heidelberg, 2012.
6. Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pages 1735– 1742, 2006.
7. Sebastian Ho¨fer, Manfred Hild, and Matthias Kubisch. Using slow feature analysis to extract behavioural manifolds related to humanoid robot postures. In 10th International Conference on Epigenetic Robotics, pages 43–50, 2010.
8. Marcus Hutter. Feature reinforcement learning: Part I: Unstructured MDPs. Journal of Artificial General Intelligence, 1:3–24, 2009.

9. Odest Chadwicke Jenkins and Maja J. Mataric´. A spatio-temporal extension to isomap nonlinear dimension reduction. In 21st International Conference on Machine Learning (ICML), page 56, 2004.
10. Nikolay Jetchev, Tobias Lang, and Marc Toussaint. Learning grounded relational symbols from continuous data for abstract reasoning. In Autonomous Learning Workshop at the IEEE International Conference on Robotics and Automation, 2013.
11. Rico Jonschkowski and Oliver Brock. Learning task- specific state representations by maximizing slowness and predictability. In 6th International Workshop on Evo lutionary and Reinforcement Learning for Autonomous Robot Systems (ERLARS), 2013.
12. Jens Kober, J. Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. International Journal of Robotics Research, 32(11):1238–1274, 2013
13. Frank Seide, Gang Li, and Dong Yu. Conversational speech transcription using context-dependent deep neural networks. In Interspeech1, pages 437–440, 2011.
14. Roger N. Shepard. Toward a universal law of generalization for psychological science. Science, 237(4820): 1317–1323, 1987.
15. Satinder P. Singh, Tommi Jaakkola, and Michael I. Jordan. Reinforcement learning with soft state aggregation. In Advances in Neural Information Processing Systems (NIPS), pages 361–368, 1995.
16. Nathan Sprague. Predictive projections. In 21st International Joint Conference on Artificial Intelligence (IJCAI), pages 1223–1229, 2009.