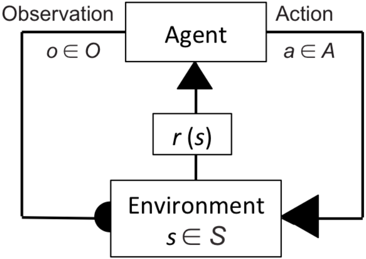
State Representation Learning in Robotics: Using Prior Knowledge about Physical Interaction

Yu-Guan Hsieh

Juin 2017

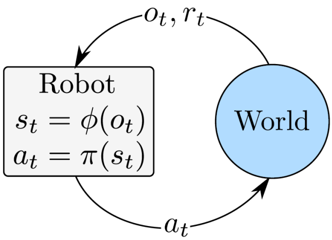
What is a robot?

- Stimilus, Action
- Flux sensorimoteur, Curiosité ...



Internal representation

- Le robot doit apprendre à agir correctement
- $\rightarrow$ établir une representation du monde réel



Prior

- Des connaisances à priori
- *Task-specific priors*: feature engineering (extraction de caractéristique)
- L'apprentissage de reprsenation: il faut utiliser *generic priors* (aussi lié à big data, calcul intensif ...)
- Plus difficile qu'une simple réduction de dimension
- *Robotic priors*: prenant en compte les lois physiques

Robotic Prior

- *Simplicity Prior.*

  Faible dimension de l'espace de representation, ce que donne à priori aussi une meilleure capcité de généralisation

- *Temporal coherence Prior.*

  Des propriétés petinentes à la tâche varient lentement.

$$L_{\text{temporal coherence}}(D, \phi) = \mathbf{E}[\|\Delta s_t\|^2]$$

- *Proportionality Prior.*

  L'effet de l'action est propotionnel à l'amplitude de l'action.

$$L_{\text{proportionality}}(D, \phi) = \mathbf{E}[(\|\Delta s_{t_2}\| - \|\Delta s_{t_1}\|)^2 | a_{t_1} = a_{t_2}]$$

Robotic Prior (Continued)

- *Causility Prior.*

  Les propriétés petinentes à la tâche et l'action ensemble détermine la récompense.

$$L_{\text{causality}}(D, \phi) = \mathbf{E}[e^{-\|s_2 - s_1\|} | a_{t_1} = a_{t_2}, r_{t_1+1} \neq r_{t_2+1}]$$

- *Repeatability Prior.*

  Les propriétés petinentes à la tâche et l'action ensemble détermine le changement en ces propriétés.

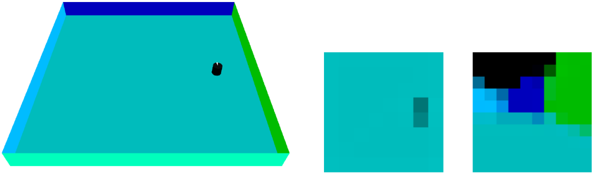
$$L_{\text{repeat}}(D, \phi) = \mathbf{E}[e^{-\|s_2 - s_1\|} \|\Delta s_{t_2} - \Delta s_{t_1}\| | a_{t_1} = a_{t_2}]$$

Technical points

- ϕ linéaire
- descente de gradient

Navigation task

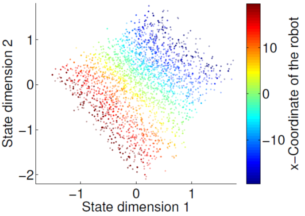
- Chambre de taille 45×45
- L'orientation fixée (vers le haut de la figure)
- Se déplacer de [-6, -3, 0, 3, 6] selon X, Y (25 actions)
- Une bruit gaussienne ajoutée à l'effet de l'action



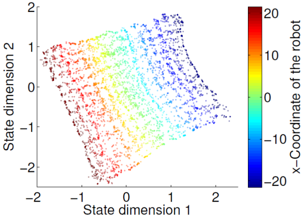
(a) Mobile robot in a square room (b) Top-down (c) Egocentric

Navigation task (Continued)

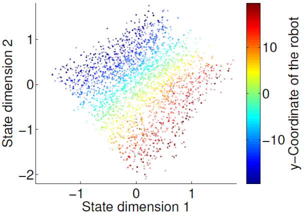
- -1 si le robot se heurte contre le mur, +10 si la distance entre le robot et le point (45,45) est inférieur à 10
- Tâche simple: vue descendante
- Tâche normale: vue égocentrique, un champ de vision de 300°
- Le robot explore l'espace aléatoirement
- L'espace de representation de dimension 2



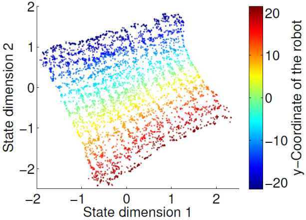
(d) Top-down view state samples (x)



(e) Egocentric state samples (x)

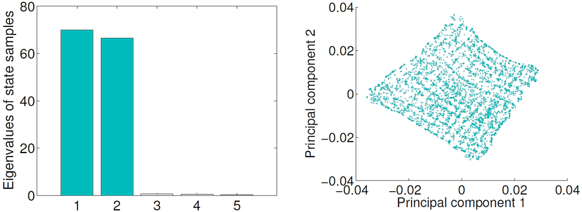


(f) Top-down view state samples (y)



(g) Egocentric state samples (y)

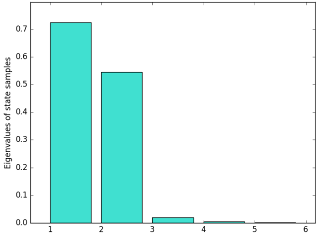
Mapping to higher-dimensional space

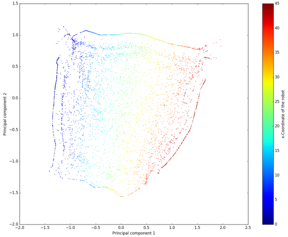


(a) Eigenvalues of state samples (b) Projected state samples

Fig. 5. Results for the navigation task with a five-dimensional state space.

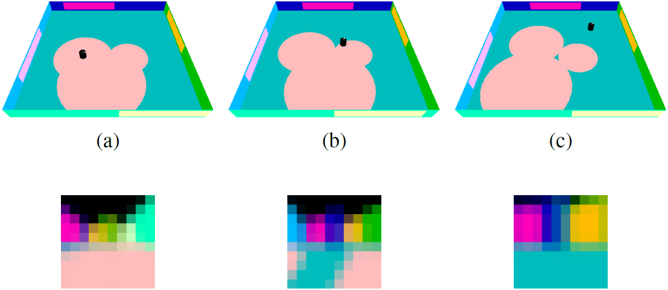
Mapping to higher-dimensional space (mine)





Extended navigation task

- Tâche plus difficile: circles et rectangles qui bougent aléatoirement, le robot peut tourner



(a) (b) (c)

(d) (e) (f)

Fig. 7. Extended navigation task. (a-c) show the robot moving to the upper right corner while the distractors move randomly. (d-f) show the respective observations (note how they are influenced by the distractors).

Put this with reinforcement learning

- Apprendre la representation et faire du Q-learning tous les 500 pas de temps, puis tester la performance
- Neural fitted Q iteration

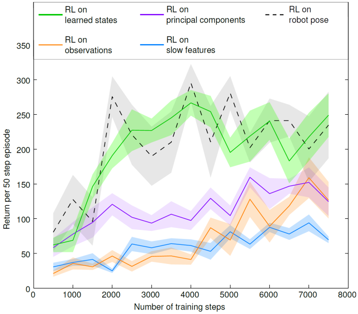
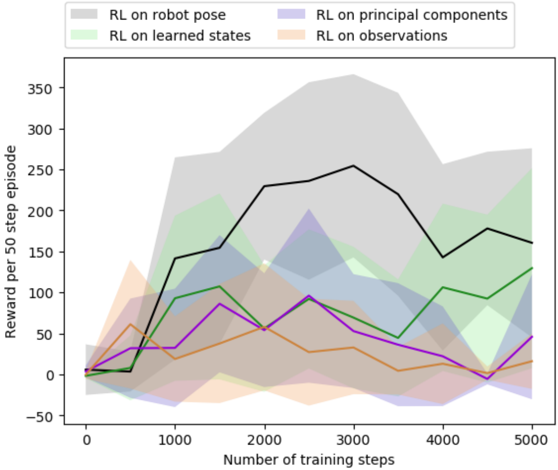


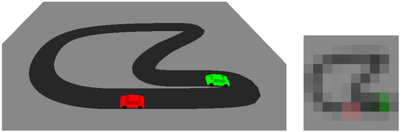
Fig. 8. Reinforcement learning performance for different state representations. Lines show means, surfaces display their standard errors.

Put this with reinforcement learning (mine)

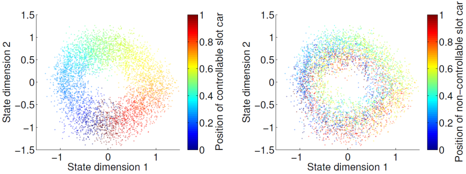


Other things

- Slot car racing task
- Ignorer les objets de distraction



(a) Slot car racing with a distractor (green car) (b) Observation



(c) State samples (red car) (d) State samples (green car)

Fig. 4. Results for the slot car racing task (a) with visual observations (b). The color relates state samples to the relevant car (c) and the distractor (d).

References

Rico Jonschkowski and Oliver Brock.
State Representation Learning in Robotics: Using Prior Knowledge about Physical Interaction.
Proceedings of Robotics: Science and Systems, 2014

Martin Riedmiller.
Neural fitted Q iteration - first experiences with a data efficient neural reinforcement learning method.
J. Gama et al. (Eds): ECML 2005. Lecture Notes in Computer Science, vol 3720. Springer, Berlin, Heidelbergi, 2005.