Making robots learn for the real world

DEVIEW 2018

Author : Tomi Silander

Team : Morgan Funtowicz, Arnaud Sors, Julien Perez & NAVER LABS robotics team

NAVER LABS

CONTENT

DEVIEW 2018

Superhuman AI – Alpha Go (Zero)

DEVIEW 2018

"Google's AI AlphaGo Is Beating Humanity At Its Own Games"

20 million self-play games

200 000 games per day

Superhuman AI – Dota2

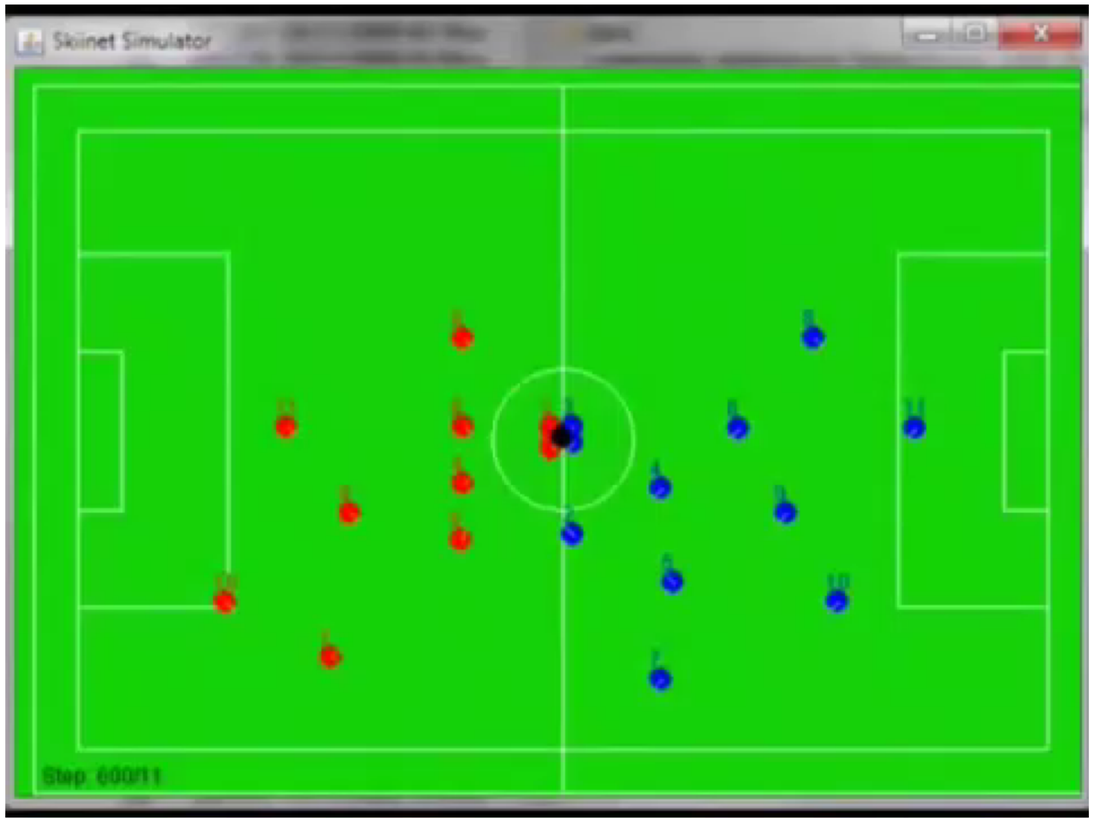
DEVIEW 2018

AI bots trained for 180 years a day to beat humans at Dota-2

Robot Soccer – subhuman AI

DEVIEW 2018

Soccer over chess as a paradigmatic robotic task (Sahota et al. AI-94)



Thus a GAP

and its been widening lately, but why?

Computational power has made it possible to scale RL i.e.,
to make the software agents able to learn end-to-end by trial and error in simulations that can be run fast and in parallel producing huge amounts of training data.

CONTENT

DEVIEW 2018

Should we make robots learn?

DEVIEW 2018

First Conference on Robot Learning 2017, keynote by Rodney Brooks:

1. Well, maybe – just for fun, to see if that is possible
2. To put correct pressure to machine learning methods - my favorite argument
3. To make robots more practical – maybe

So if the Godfather of robotics is not too enthusiastic, we might also want to be skeptical

Learning in robotics is a hot topic

DEVIEW 2018

1. Conference on Robot Learning 2018 (CORL)
2. Robotics: Science and Systems 2018 (RSS)
   - learning to grasp
   - learning to localize
   - transferring from simulations to real life

3. International Conference on Robotics and Automation 2018
   - lot of deep learning

4. International conference on Intelligent Robots and Systems
   - last week in Madrid

Reinforcement learning

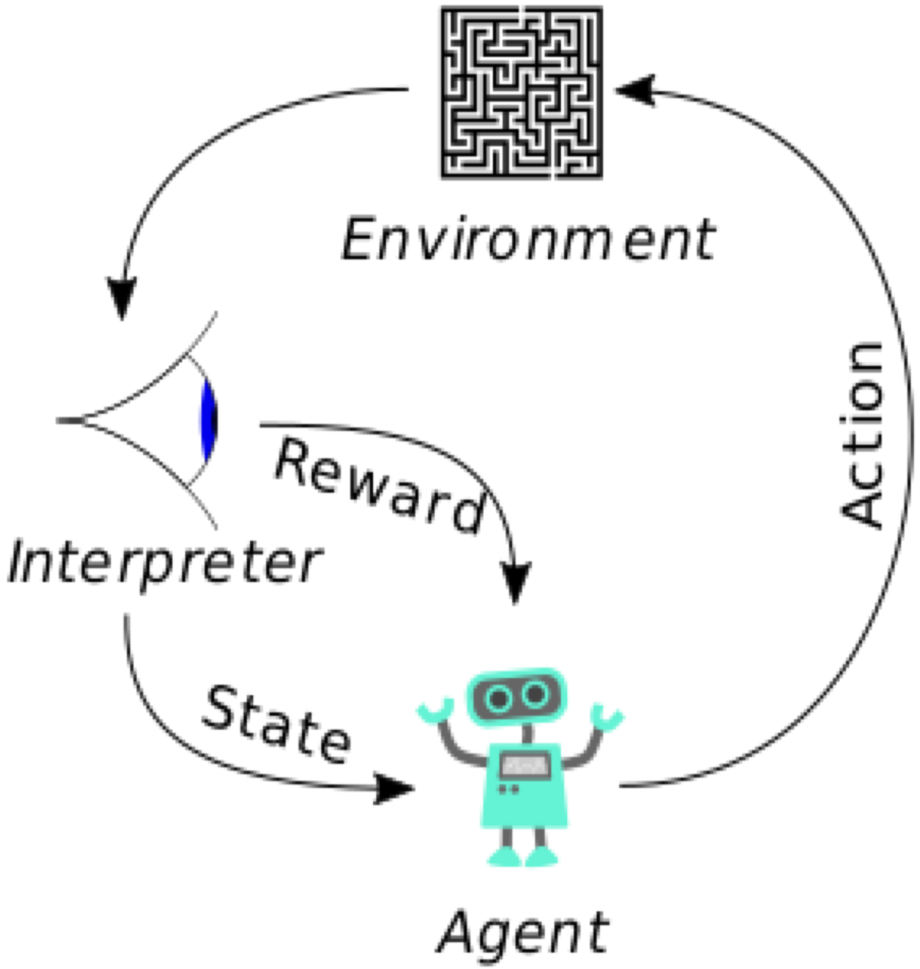
Improving behaviour using evaluative feedback via trial and error.

Target is to learn a behavioural policy:

- "In this situation S it is usually best to do action A in order to accomplish the task"
- accomplishing the task is signaled giving agent a reward for achieving the goal.

The agent learns to optimize it behaviour to get maximal reward

- all this can be formalized using decision processes.



Problems when using RL in real robots

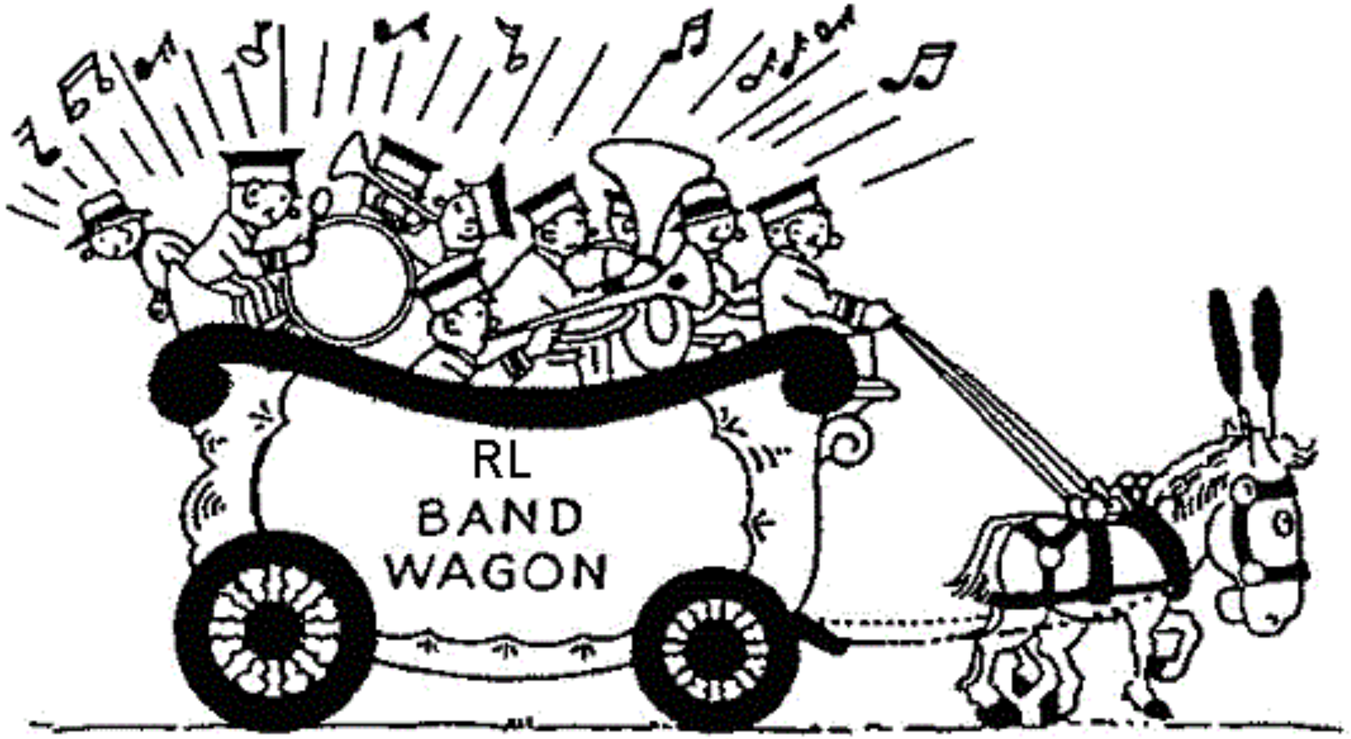
DEVIEW 2018

1. Trying all kind of things is dangerous.
   - it breaks robots and robot lab.
2. Real robot cannot be run in hyper-speed
   - so it will take "forever" before the robot learns.
3. Many tasks are not game–like
   - so robots can not play against each other.
4. Reward signals hard to give.
   - this often requires human interaction.

From Yann LeCun *How Could Machines Learn as Efficeintly as Animals and Humans?*

Before we jump on the RL bandwagon

DEVIEW 2018



other type of robotic learning
(not by trial and error)

Instruction

DEVIEW 2018

Josh Tenenbaum in his ICML keynote 2018:

"after 18 months, human children mostly learn via language"

Should we make robots understand us so we can just tell them what to do?

- but Tenenbaum talked about instruction from human to human.
- by 18 months old, children have enough "common ground" for instruction by "being told".

Thomas Nagel in The Philosophical Review 1974:

- "What is it like to be a bat?"
- robots "lifeworld" is so different that language communication will be difficult.
- many competencies (e.g., grasping) are also difficult to "explain".

Imitation learning: mimic the expert

DEVIEW 2018

One of the favorite methods in robotics, because of its safety and efficiency!

How:

- supervised learning from state to actions (lot of algorithmic variants)

Why not:

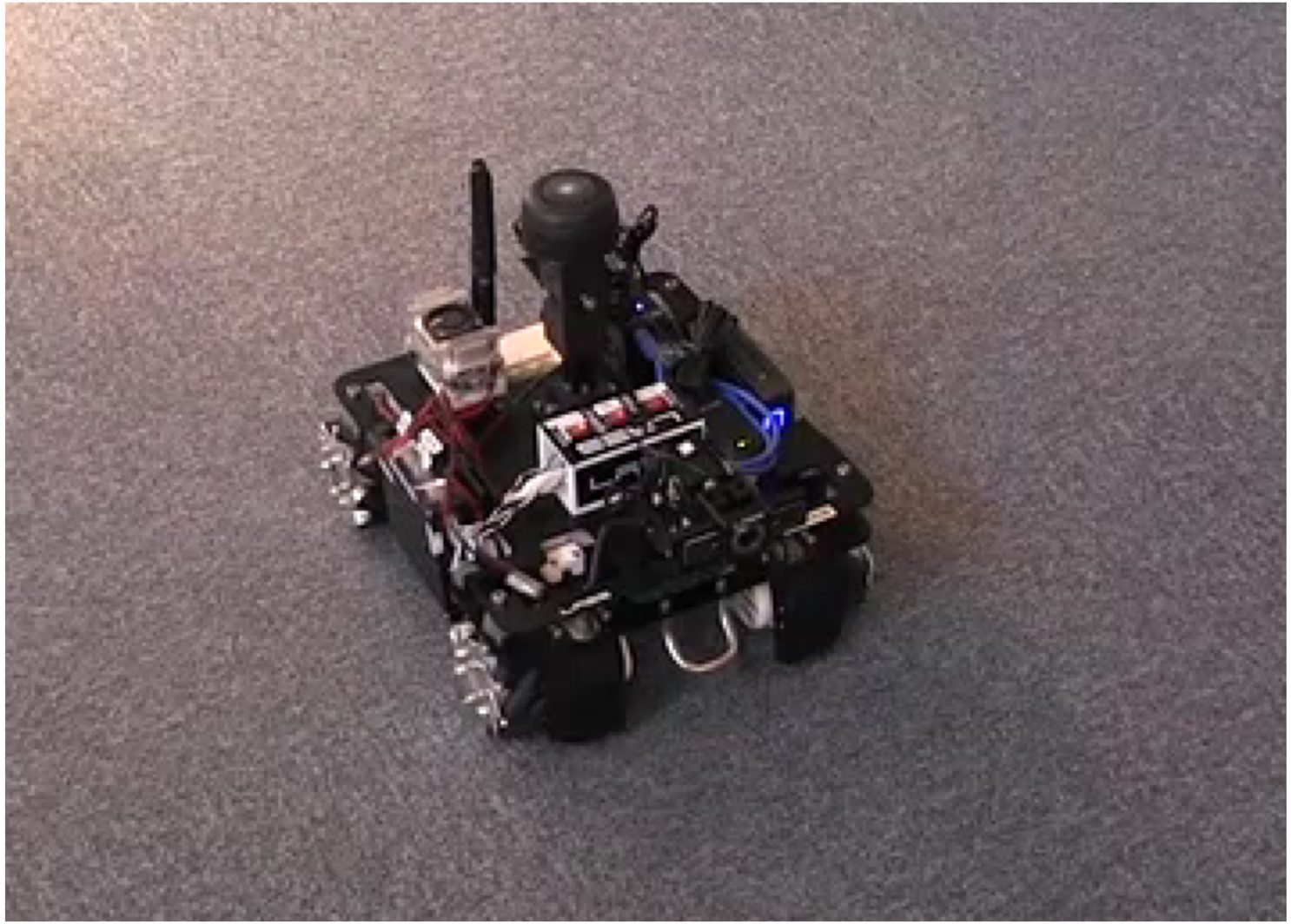
- expert cannot demonstrate behaviour on all possible situations
- so what to do in new situations?

Why traditionally a favorite method:

- robots have been used in controlled environments.

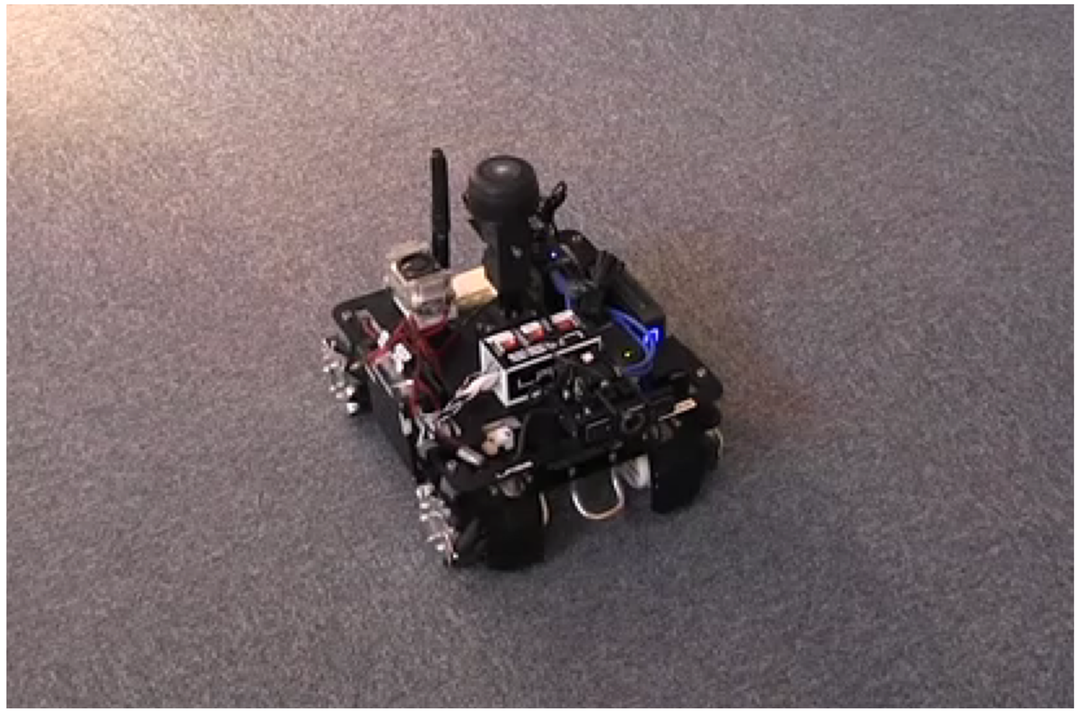
If surprise then HALT!

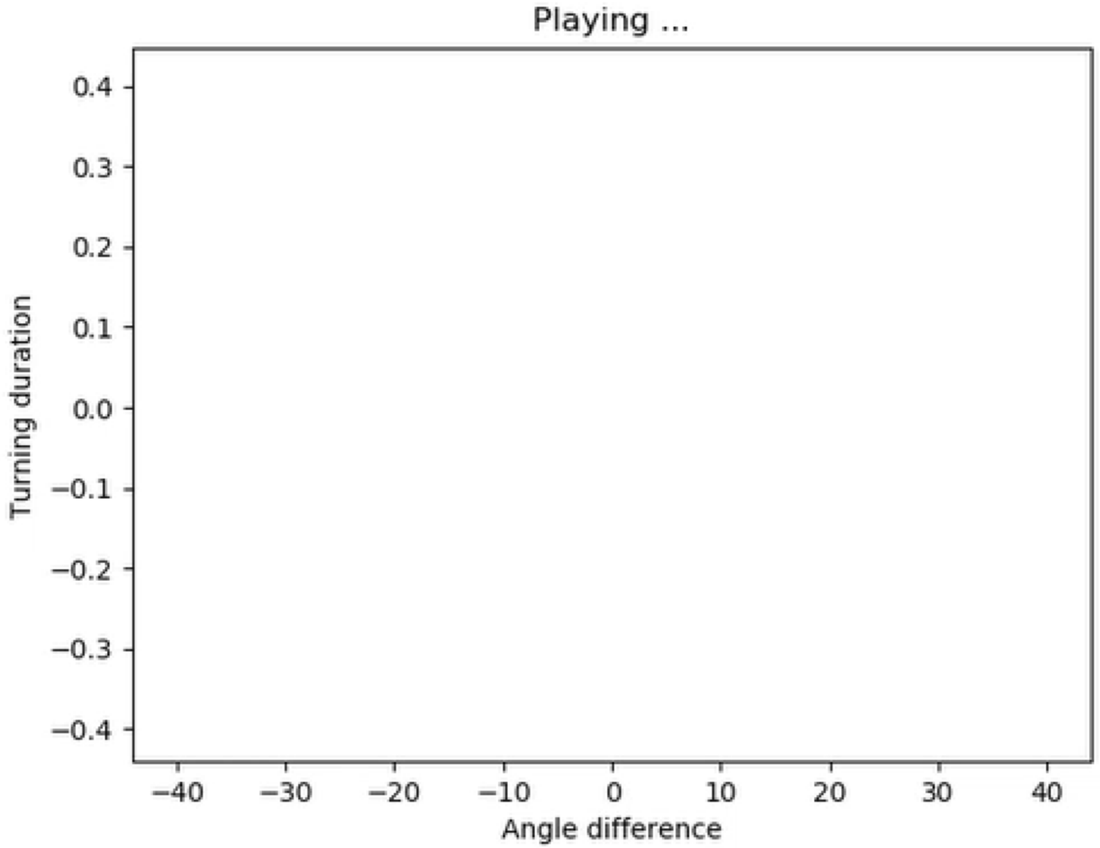
Playing around: trial, no error



Playing around, recording what happened

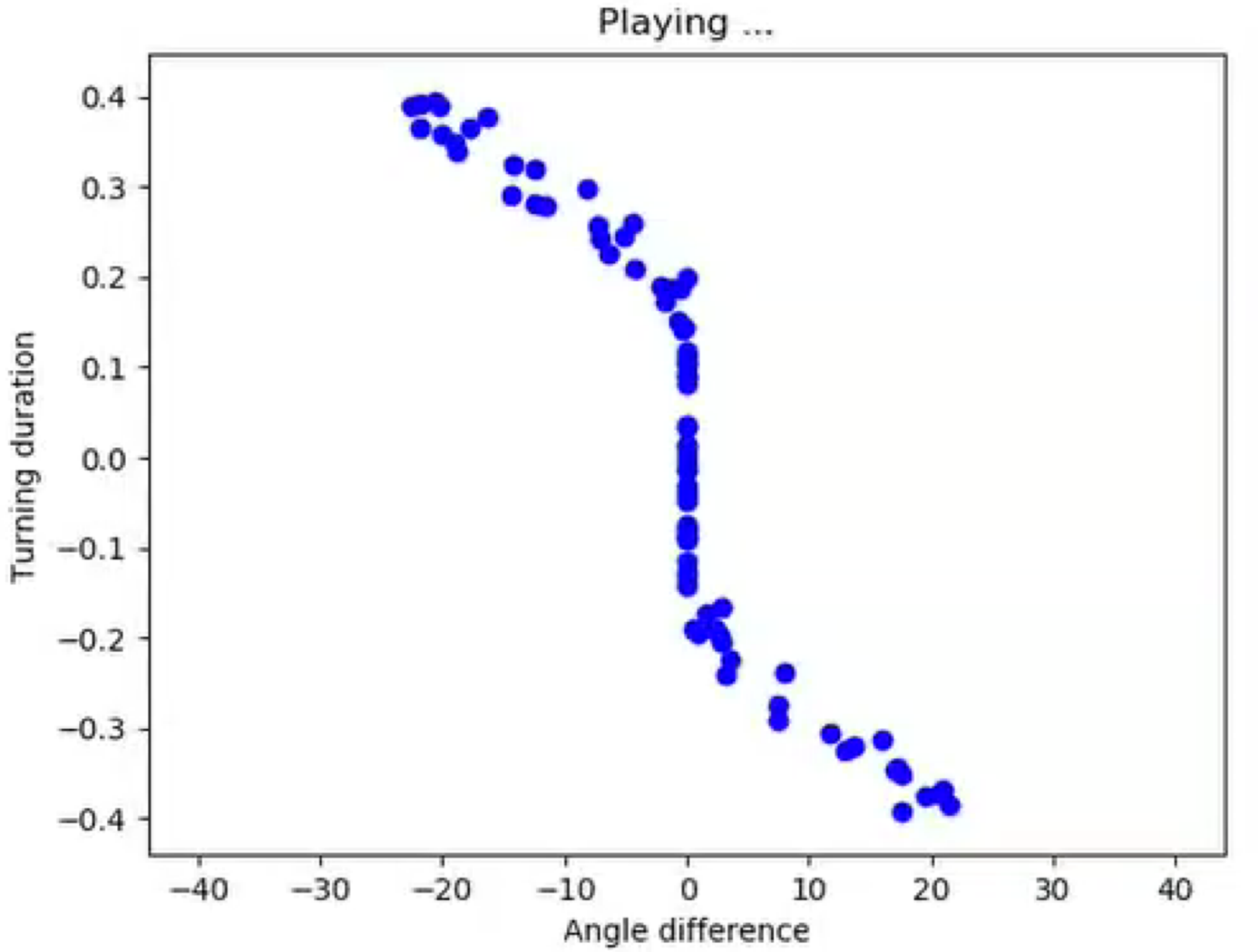
DEVIEW 2018





From playing to goal oriented actions

DEVIEW 2018



Collecting (s_t, a_t, s_{t+1}) allows us to train the regression/classification model:

$$g(s_t, s_{t+1}) = a_t.$$

If I am in situation s_t, and I want to get to situation s_{t+1}, I should do action a_t.

Self-supervised learning

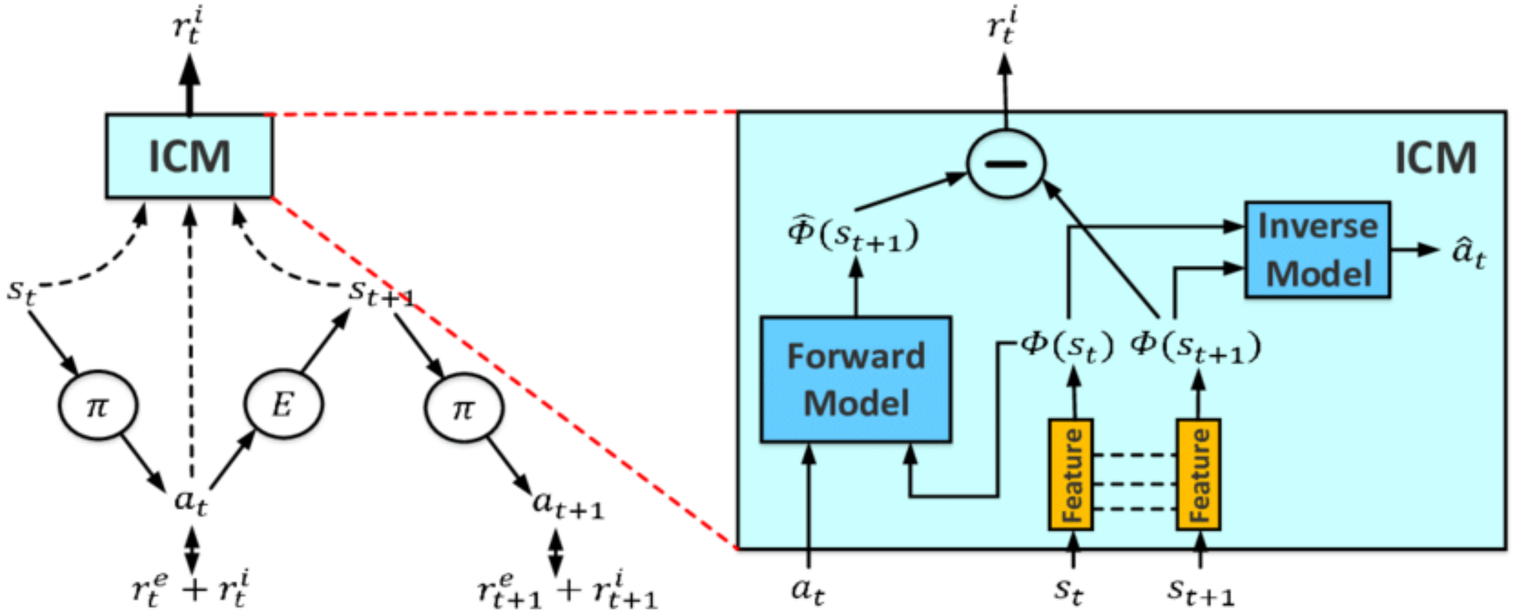
DEVIEW 2018

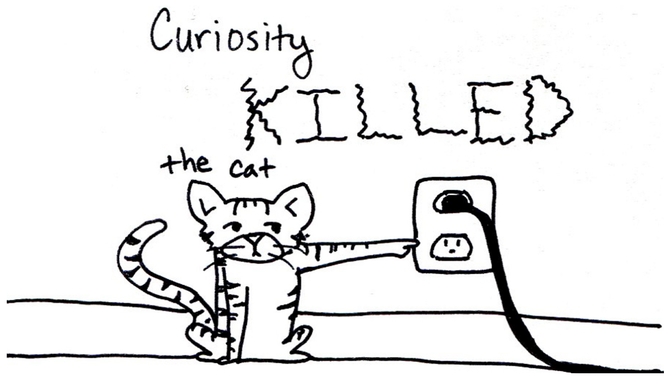
Learning general tasks that are useful for many kind of other tasks:

- grasping, obstacle avoidance, etc.

Curiosity (Pathak et al. ICML 2017):

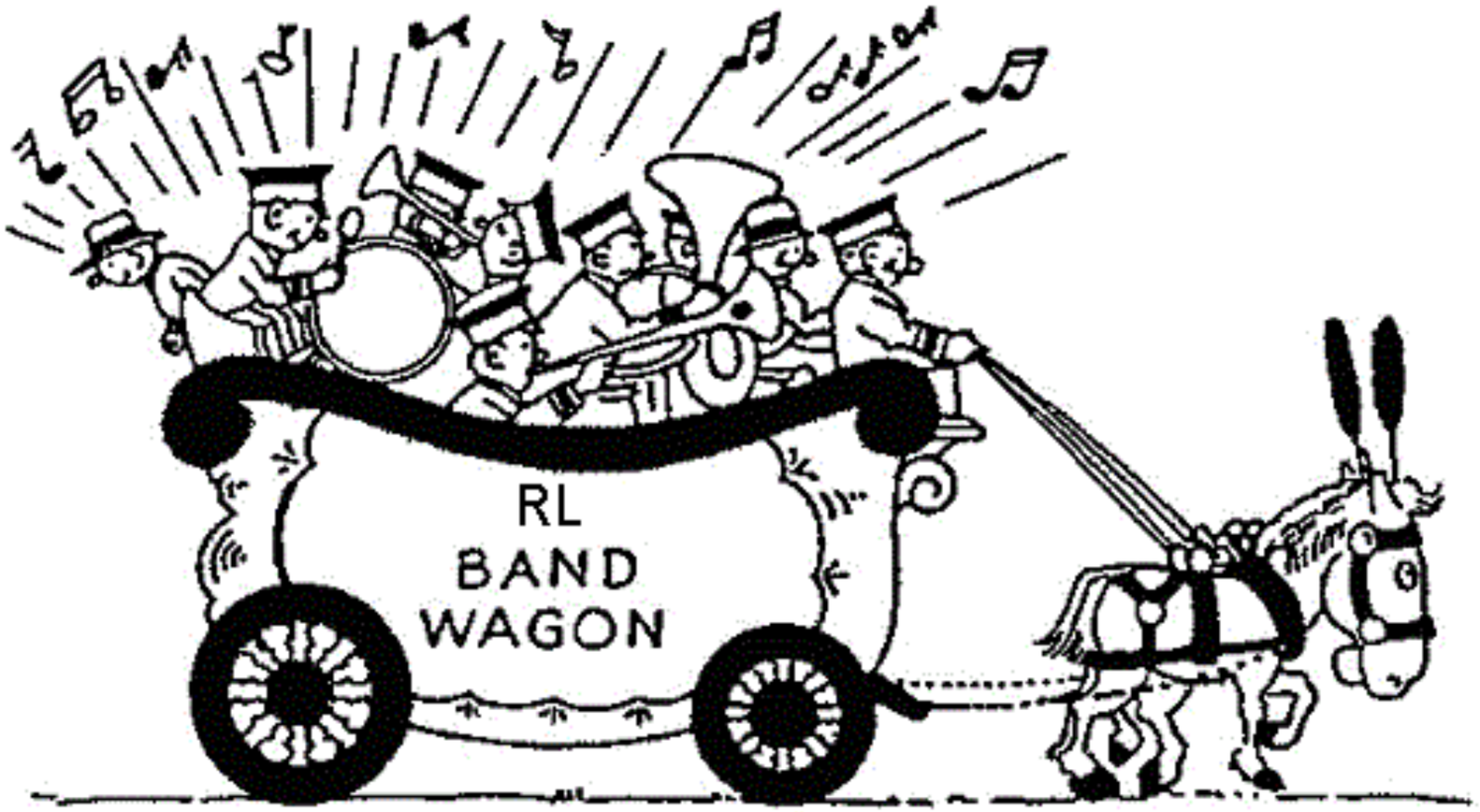
- trying things consequences of which we cannot yet predict
- learning to explore efficiently



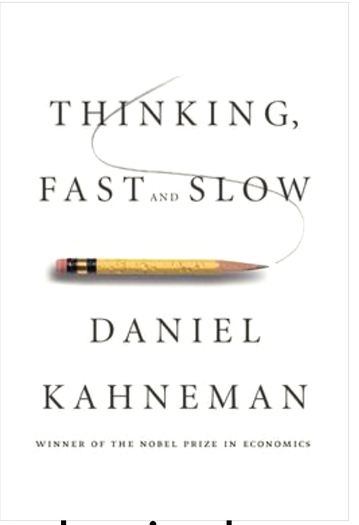


So if we really jump on RL bandwagon

DEVIEW 2018



Thinking, Fast and Slow



DEVIEW 2018

System 1:

Fast, automatic, frequent, unconscious

- see if an object is further than another
- localize the source of a sound
- display disgust when seeing a gruesome image
- drive a car on an empty road

System 2:

Slow, infrequent, logical, conscious

- look out for a woman with the gray hear
- dig into your memory to recognize a song
- determine the appropriateness of a behavior in a social context
- count a number of A's in a certain text

Similar dichotomy in RL: System 1 = model-free RL, System 2 = model-based RL

Model-free RL for thinking fast

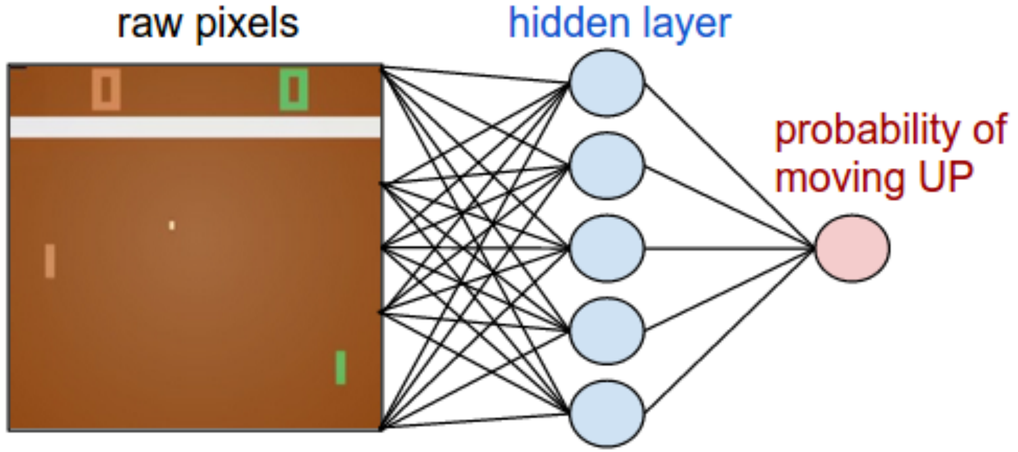
Learn the gut-feeling:

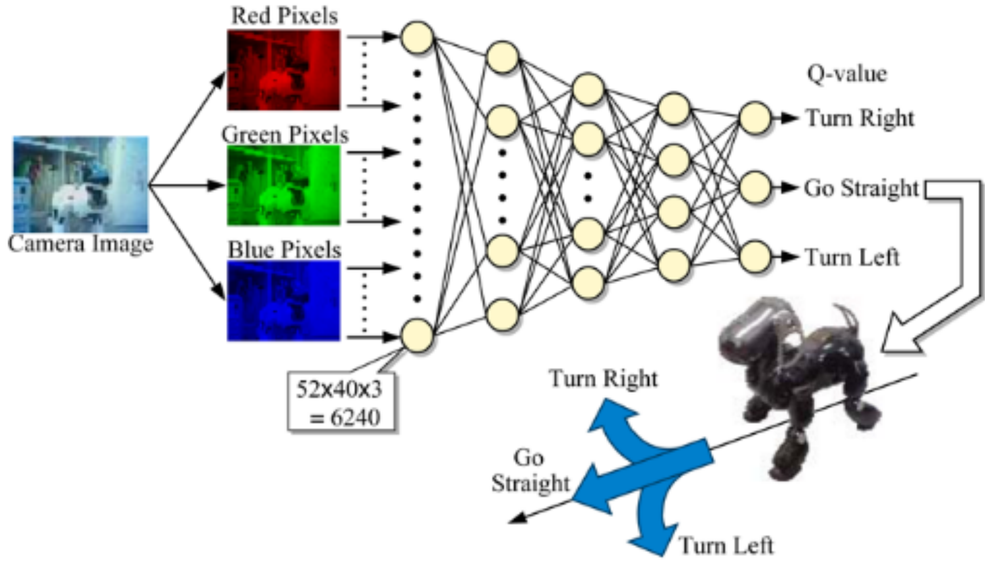
- search in policy space to learn $a(s) = \pi(s;\theta)$.
- after fitting parameters θ^*, just do $\pi(s;\theta^*)$.
- can use expert examples

Suitable when task is "simple"

Deep Q-learning (DQN):

- learn to estimate expected reward of doing action *a* in situation *s*, i.e. $Q(s,a)$.
- in each s just pick $\underset{a}{argmax}\, Q(s,a)$.





Model-based RL for thinking slow

DEVIEW 2018

Learn the model of the world, i.e. consequences of your actions:

- a "forward model" $T(s, a, s')$, i.e.,
- the probability that action *a* changes the situation *s* to *s'*.

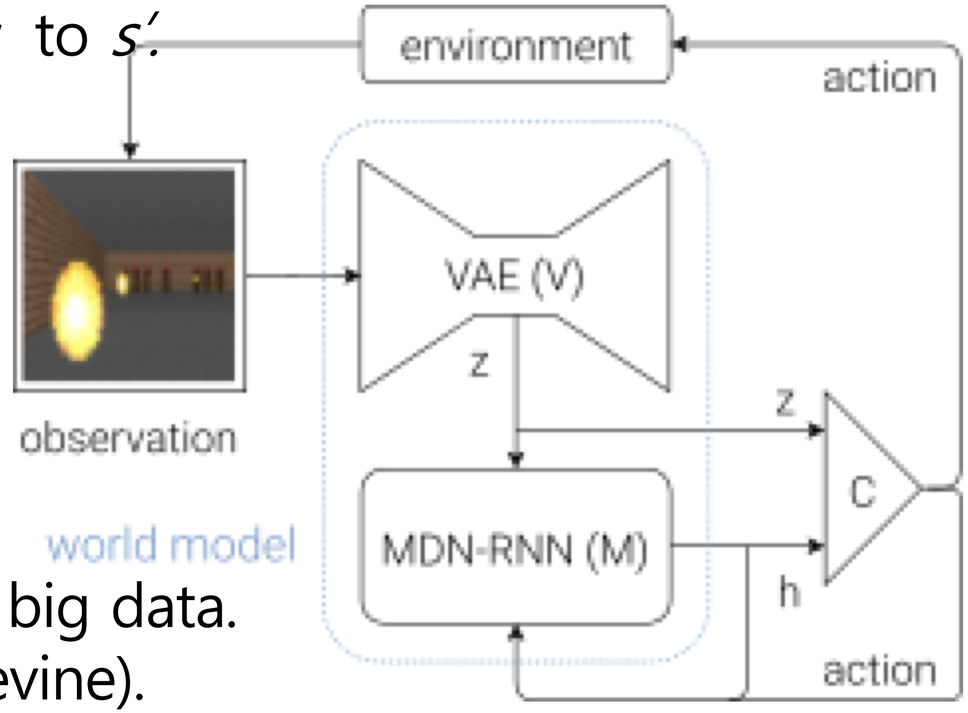
Same model can be used for many tasks.

Schmidhuber (2018):

- "World Models" => policies are simple

Dreaming:

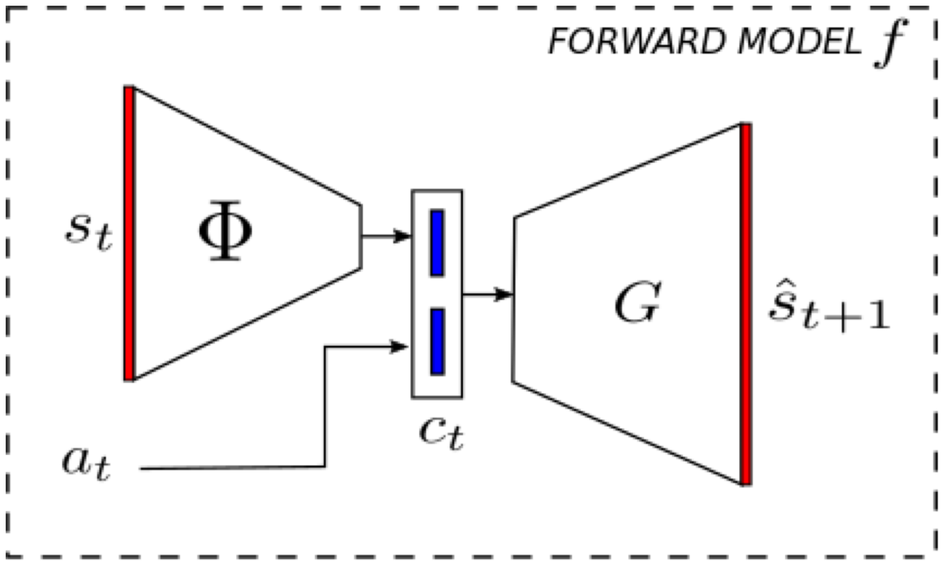
- forward models can be used to simulate the world.
- in hyper-speed, thus allowing methods that require big data.
- not realistic, but it's the versatility that matters (S. Levine).

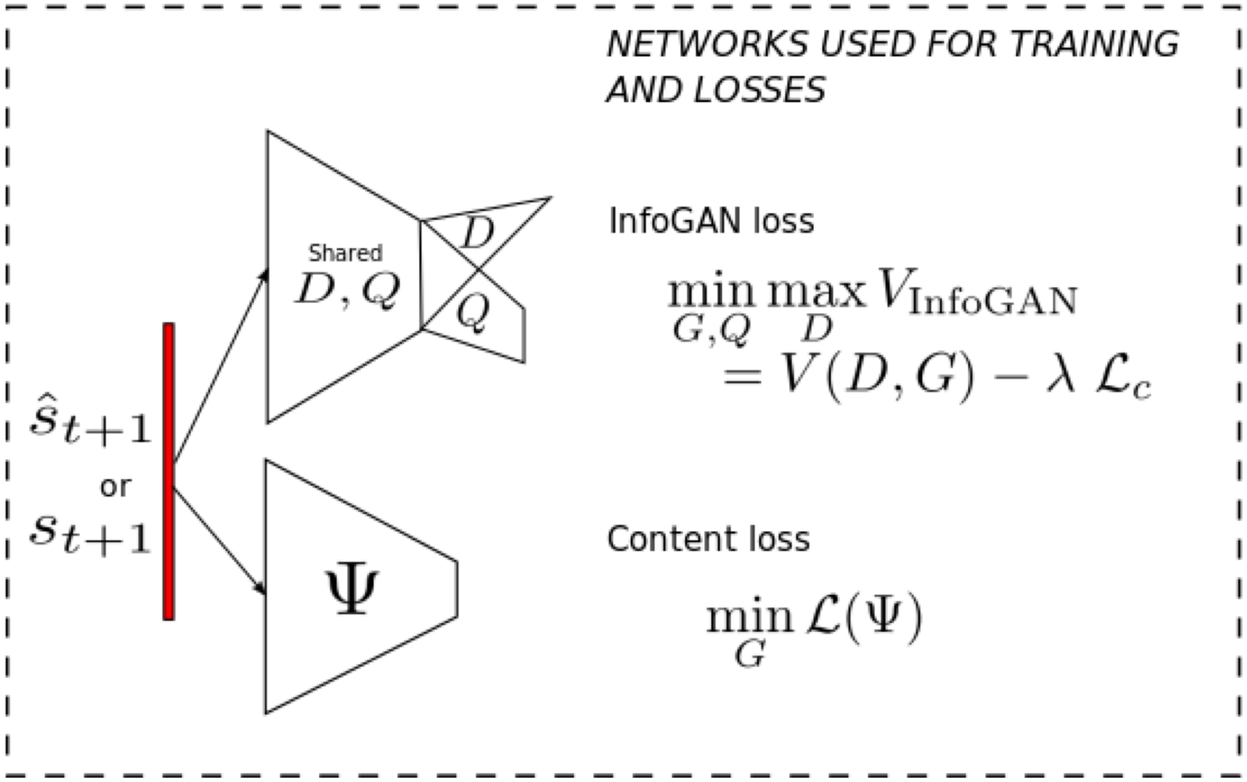


Our adversarially trained forward model

DEVIEW 2018

Combines adversarial training and ideas from Info GANs

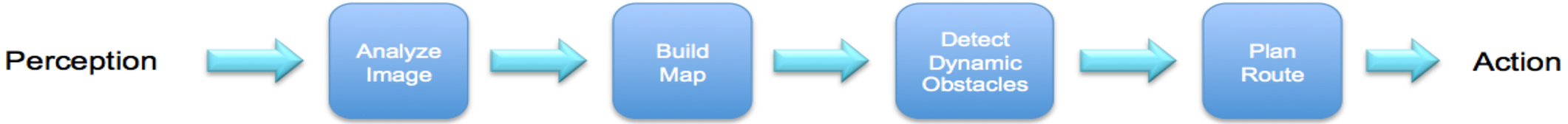




Model based or model-free RL?

NB! People have **both** System 1 and System 2

- rarely done in current RL, but maybe a key to robotics, since



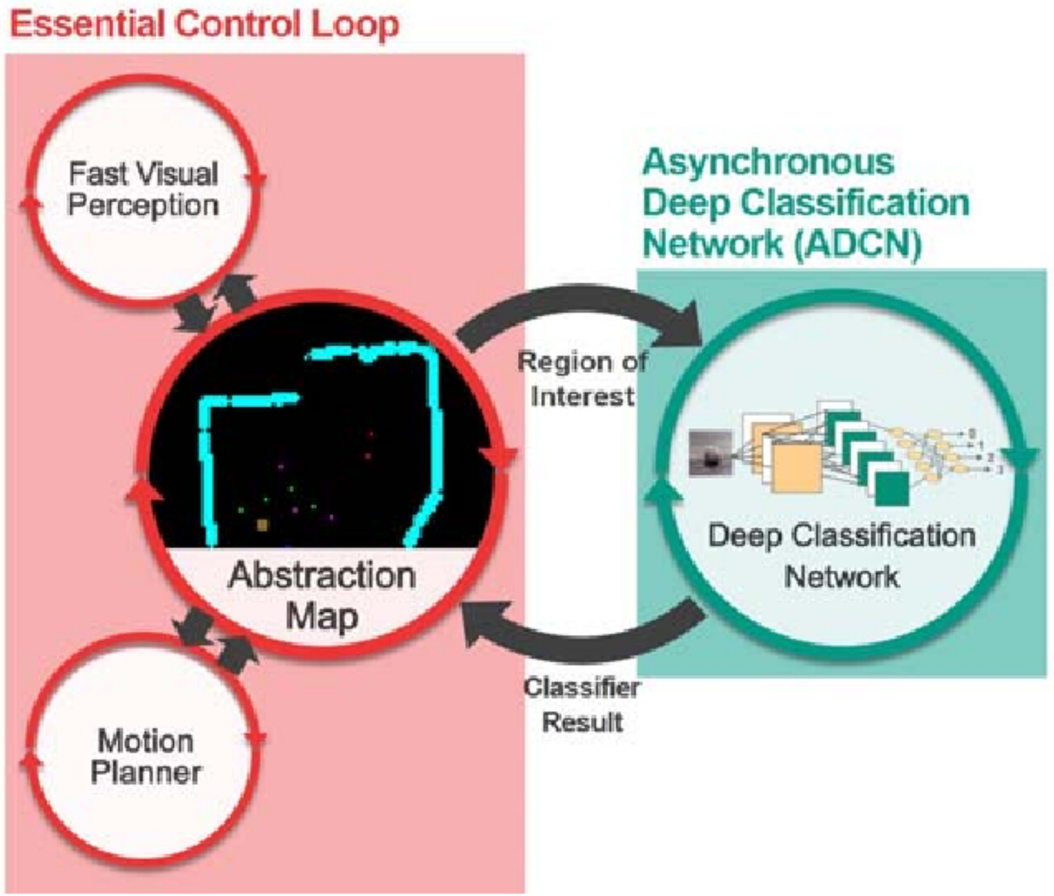
versus

Example of closing the gap

Cognitive architecture for a robot

Processes working asynchronously at different speeds:

- fast for motor control
- slower for object classification
- connected by a central representation that can also be trained via simulation



Gilhyun Ryou, Youngwoo Sim, Seong Ho Yeon and Sangok Seok (ICRA 2018)

Applying Asynchronous Deep Classification Networks and Gaming Reinforcement Learning-Based Motion Planners to Mobile Robots

Thus enter the simulations

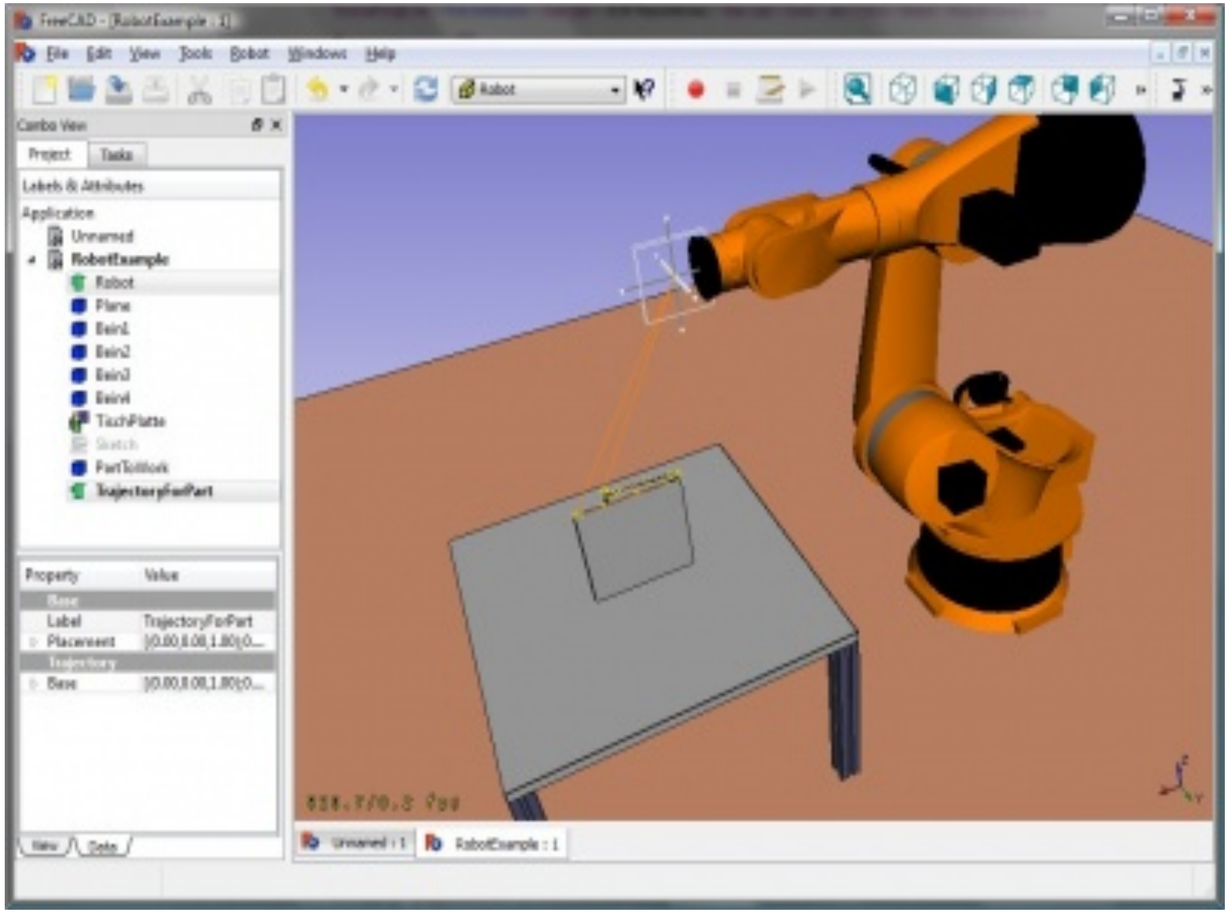
DEVIEW 2018

Huge bulk of current “robotic” RL is conducted simulations (it’s faster, safer)

- and then, sometimes(!), tried to transfer to a real robot
- maybe with some *domain adaptation*

Works OK for

- simple spaces and controlled environments where similarity of simulation and reality is easier to establish



CONTENT

DEVIEW 2018

Active Neural Localization

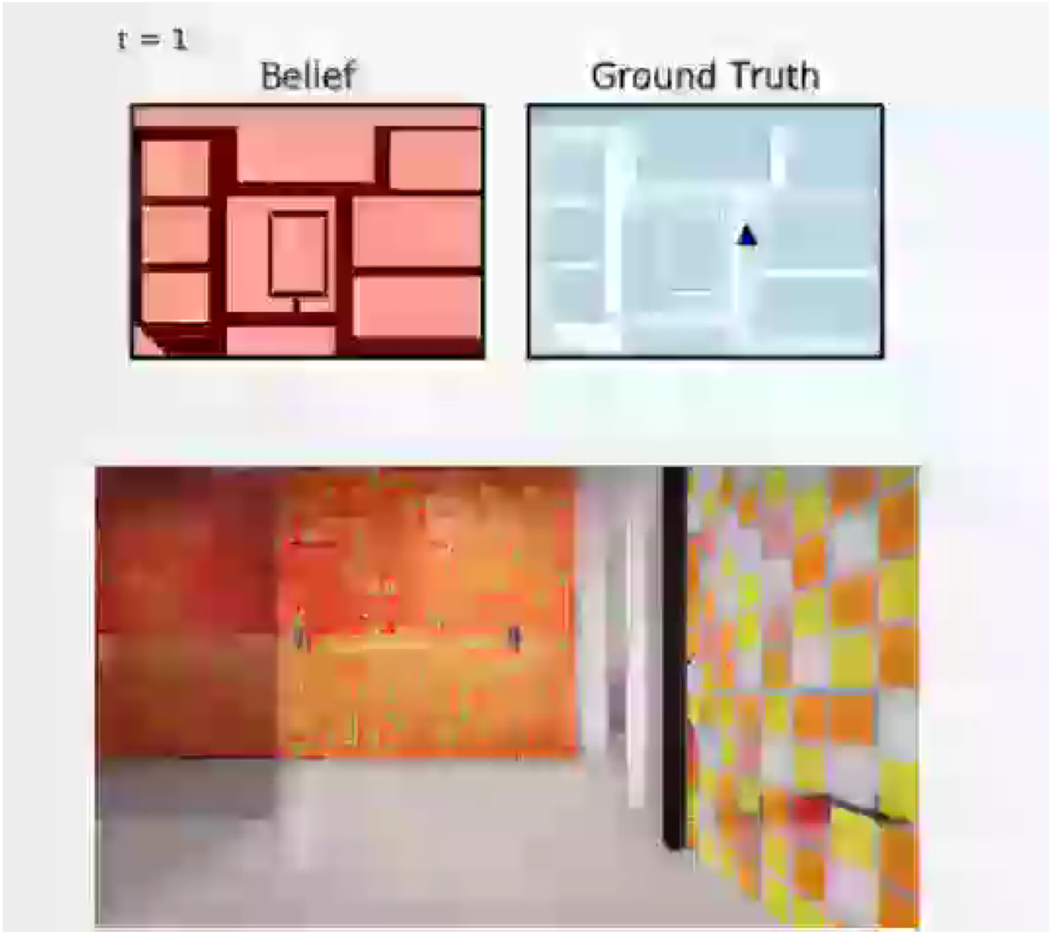
DEVIEW 2018

Perception is active, information seeking action (Gibson 1976)

If you are lost, try to look around

D.S. Chaplot et al. (NIPS 2017)

- grid presentations
- discrete angles
- noiseless operations
  - noiseless directions
  - noiseless images

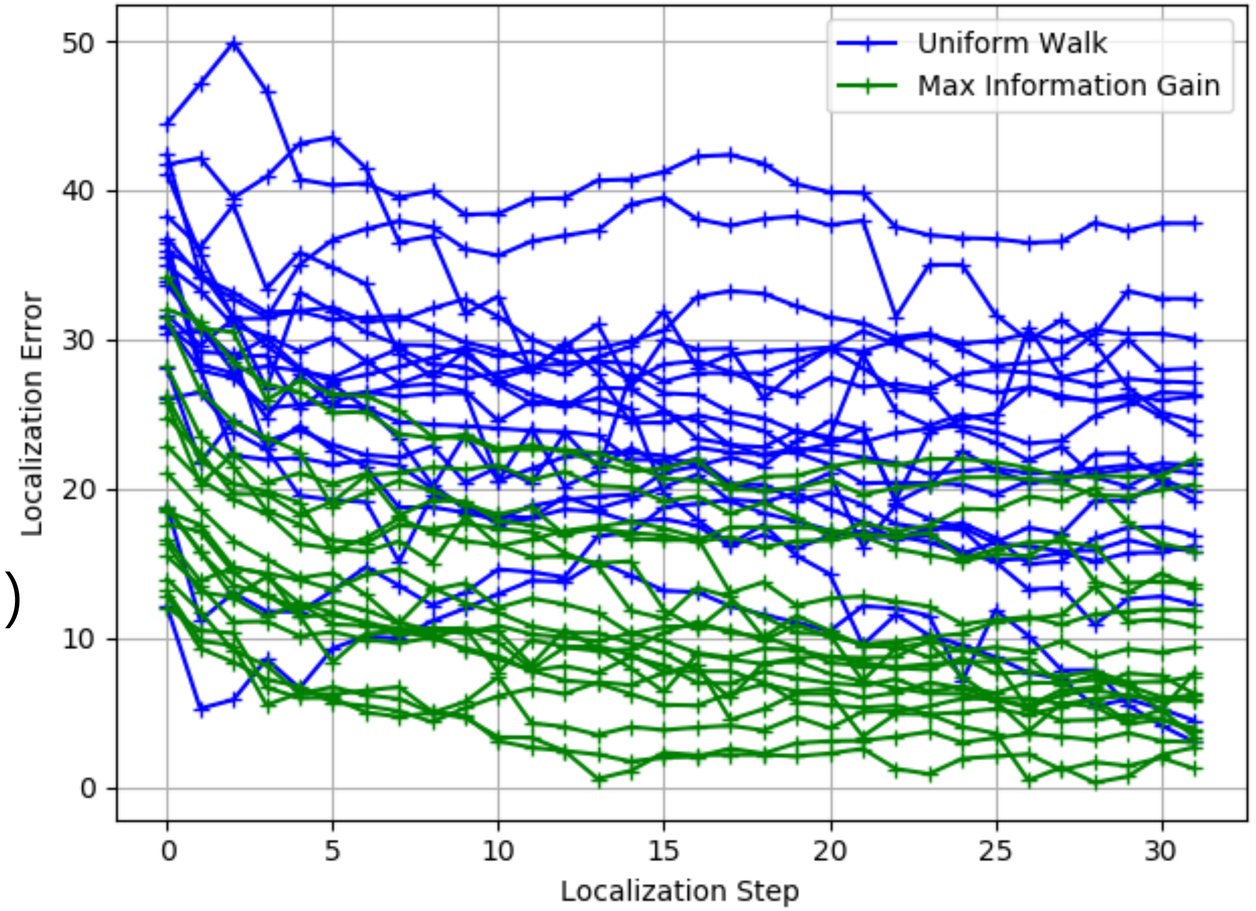


In active localization project we had to

DEVIEW 2018

1. Use particle filters instead of grid
2. Continuous, noisy directions
3. Noisy moves
4. Uncertainty in observations

After all that, yes, an active policy (still) performs better than a passive one!



Summary

DEVIEW 2018

There is a big GAP

Community is working hard to close it

- but no evidence yet for the gap diminishing

Some of the promising ways to diminish the gap:

- let the robot play to find out how the world works
- build the simulations to match the reality, not the other way round
- let the robot dream to play in alternative realities.

Q & A

질문은 Slido에 남겨주세요.

sli.do

#deview

TRACK 4

Thank you