An Investigation through Knowledge Sharing with Dynamic Networks in Reinforcement Learning

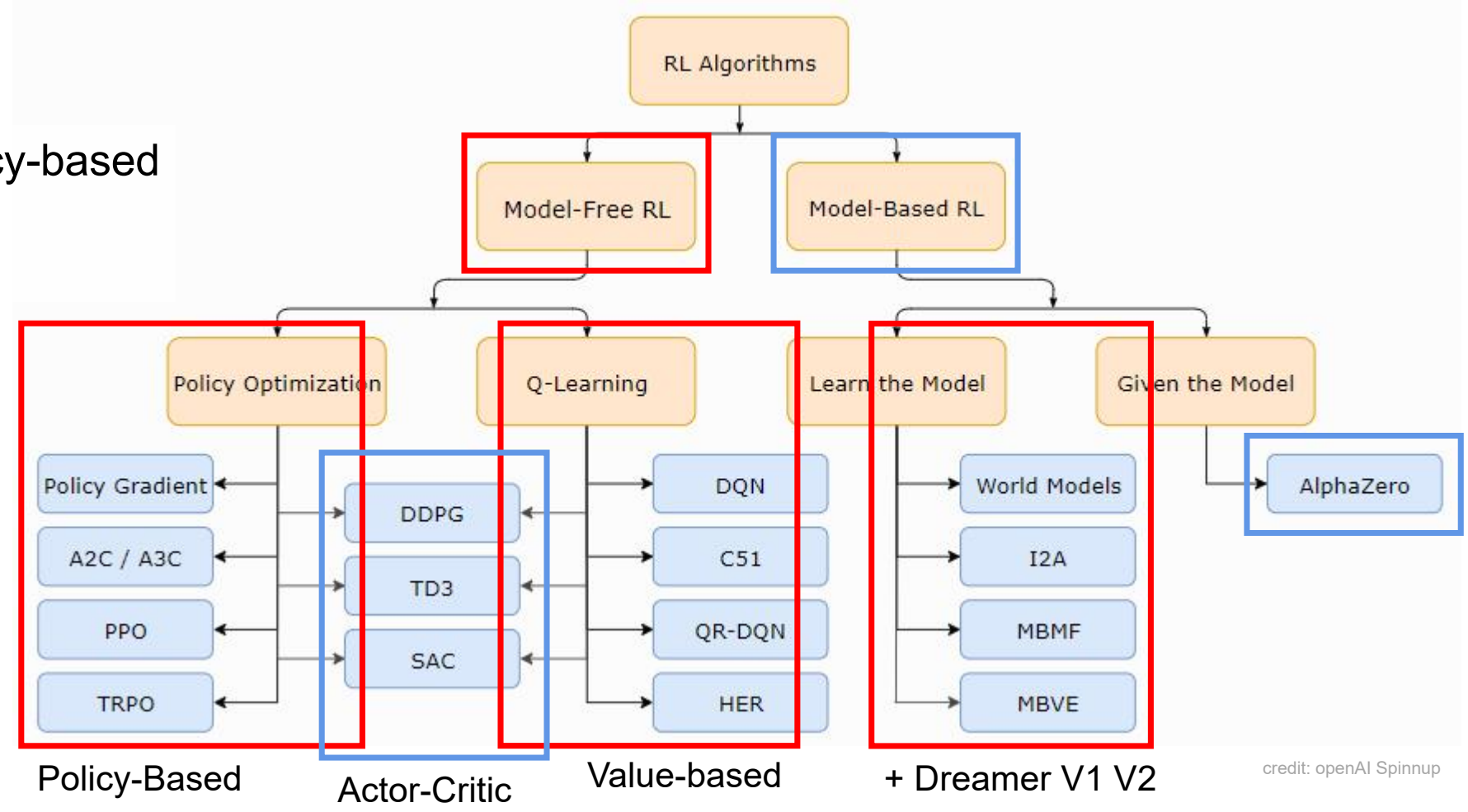
Is Exponential Accelerated RL possible?

AIR · Chengyu Jiang

Progress by the RL community

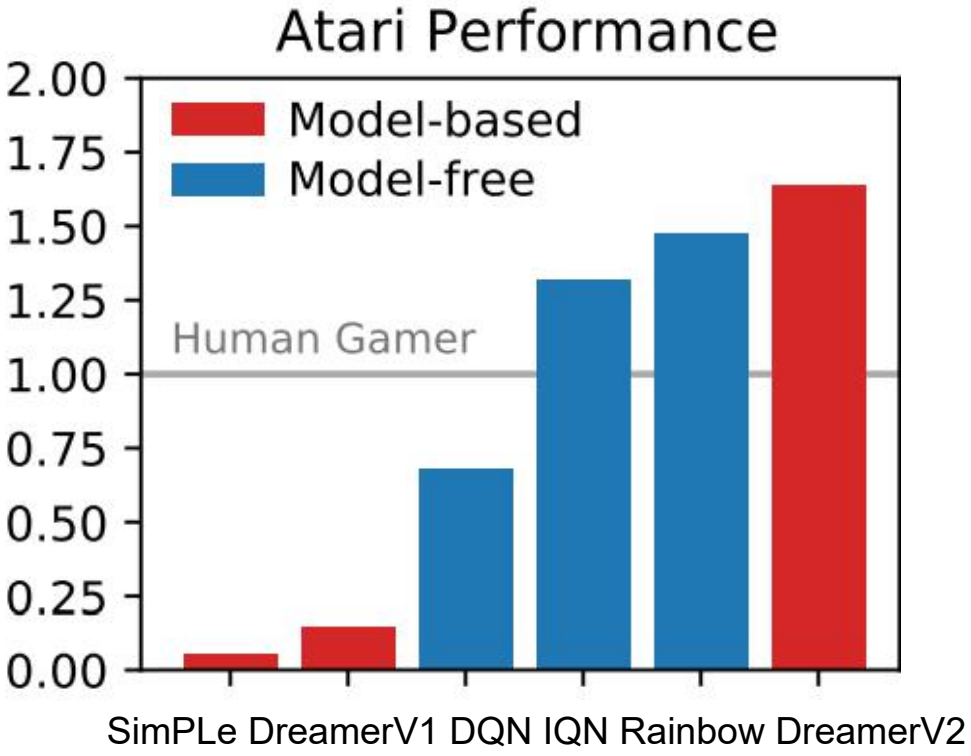
Value-based vs. Policy-based

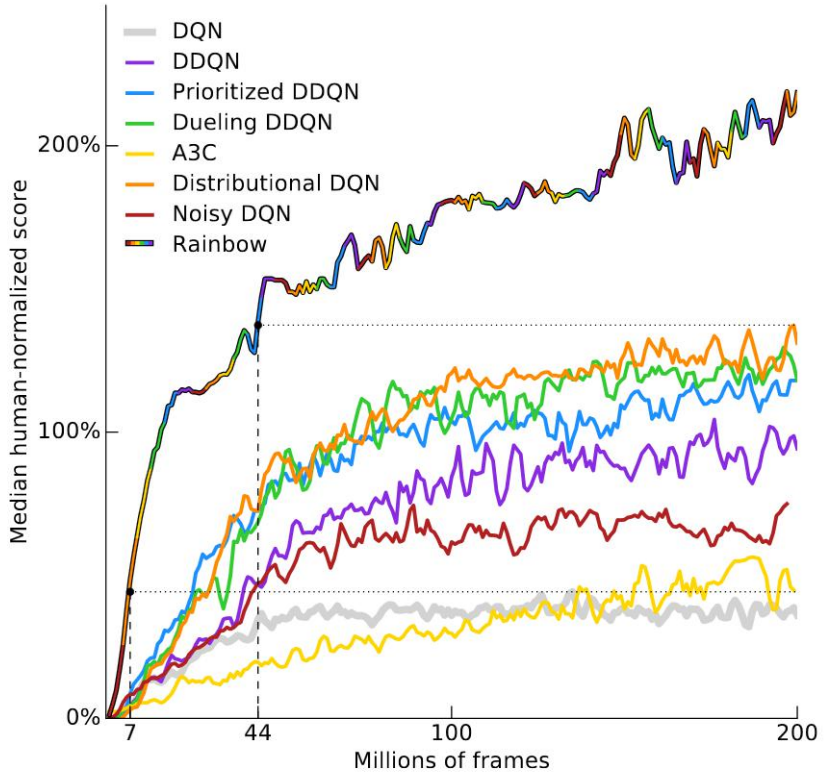
MBRL vs. MFRL



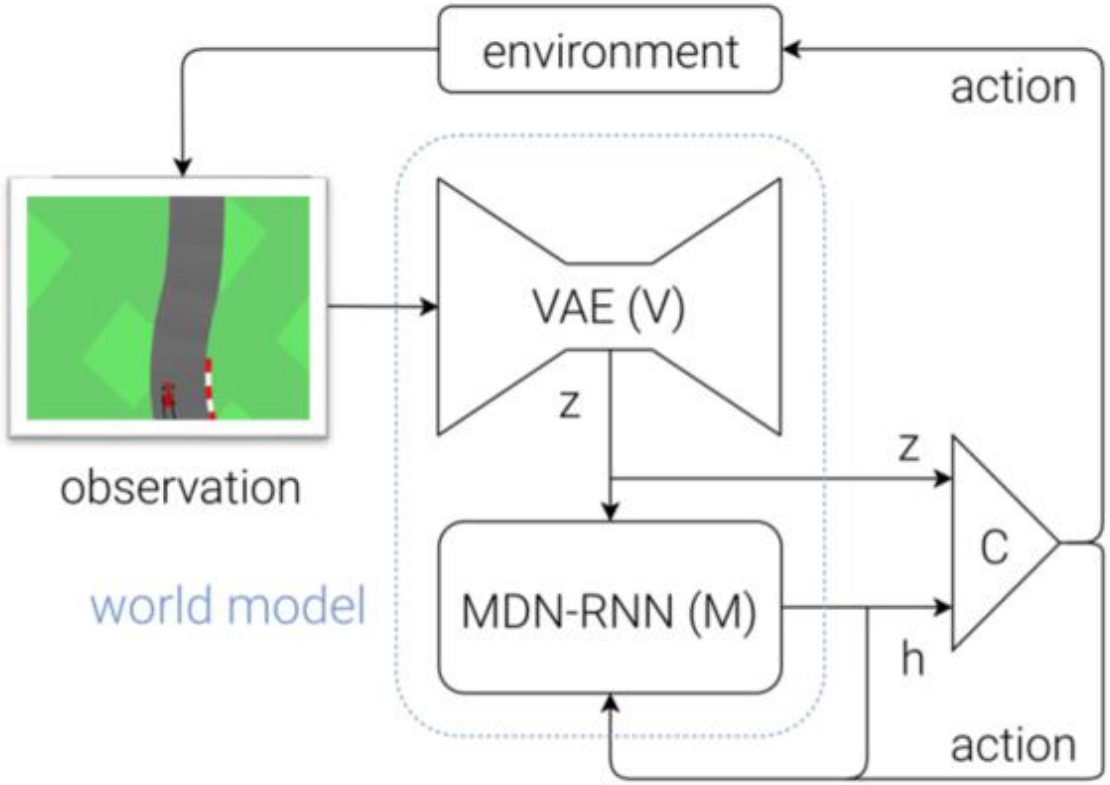
(Offline doesn't mentioned here)

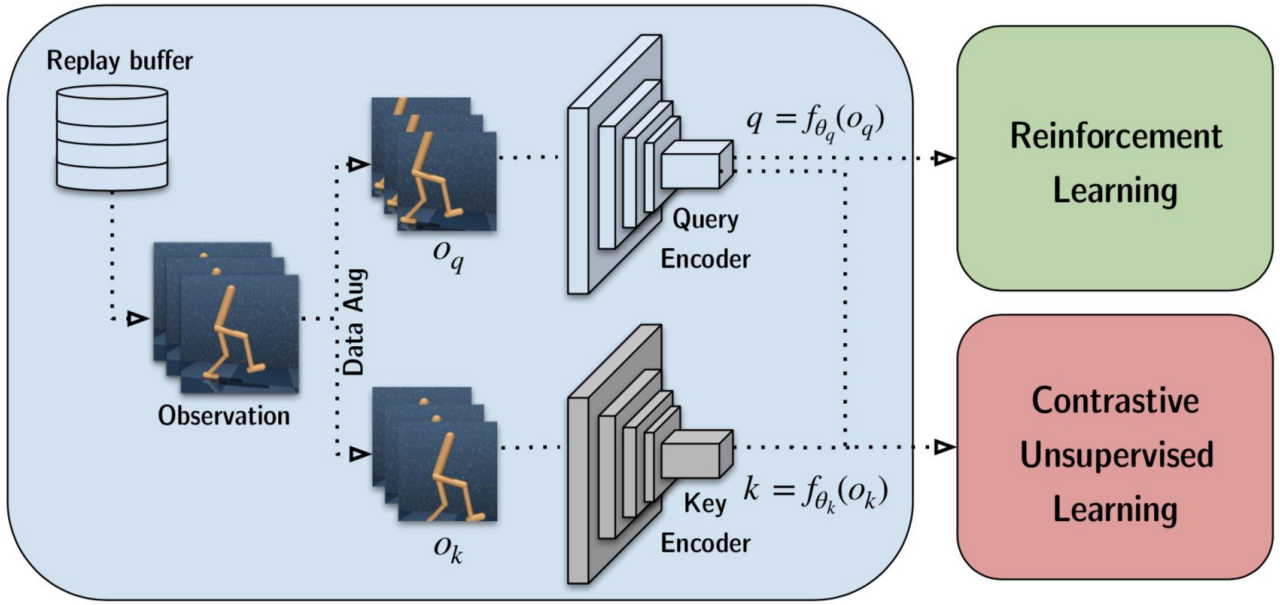
Data Efficacy Limited the Application
while the Improvements get saturated





DQN	DDQN	Dualing DQN	Distributional	Rainbow	WorldModels	Dreamer	DreamerV2
2014	2015	2016	2017	2018	2019	2020	2021



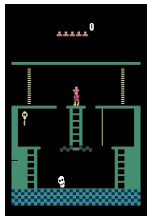


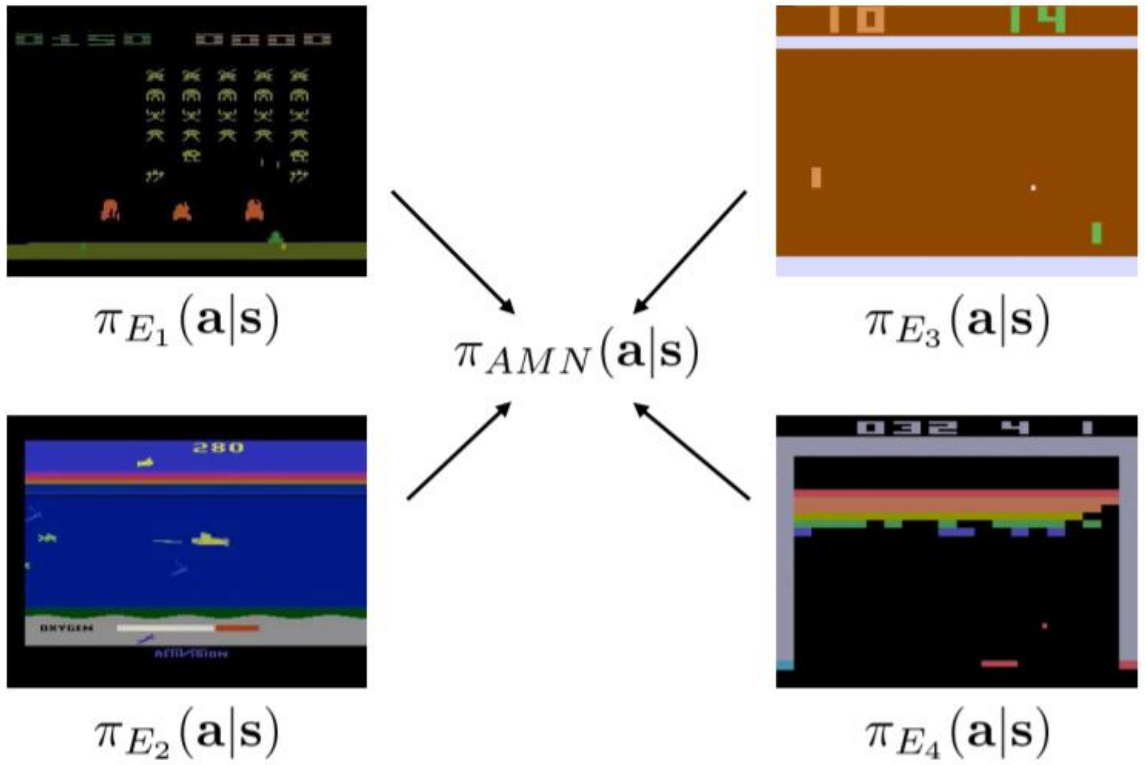
Solution to the cons, contemporarily/nowaday:

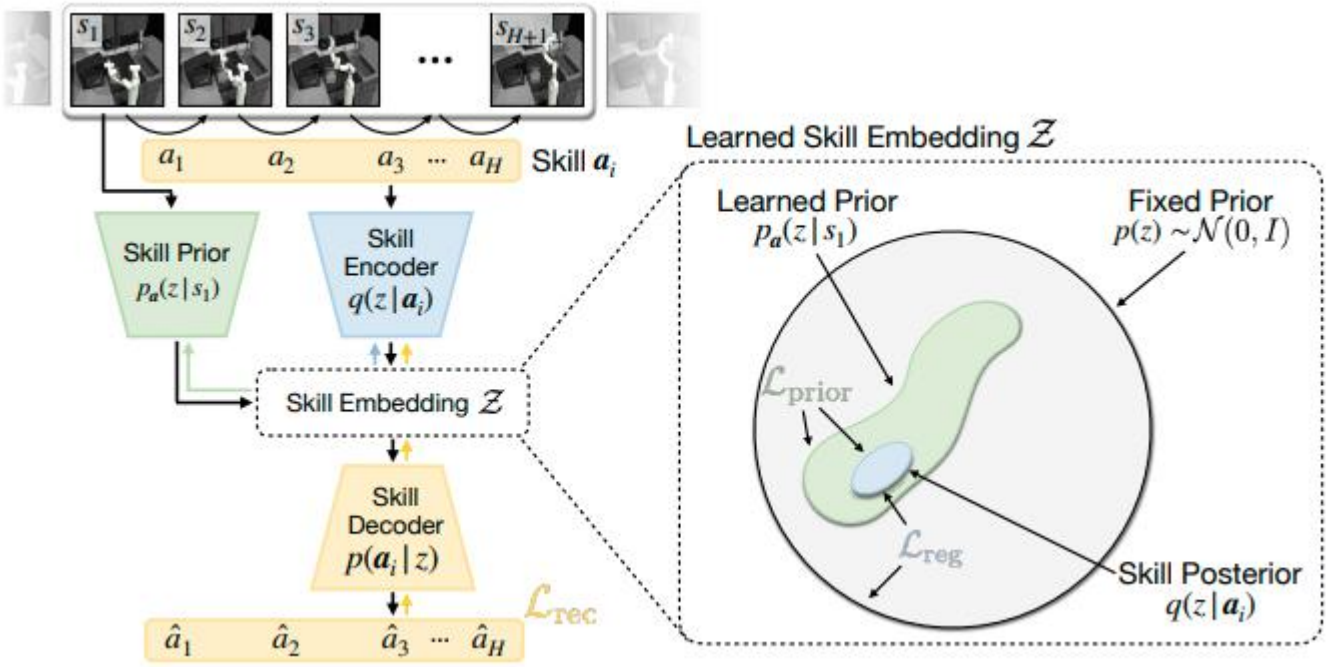
Self-modeling: World Models
Data Augmentation: Contrastive learning on visual representation (and some RL representation

x30+ data efficiency

Single Agent Plays Single Game







beyond that, what we gonna do (the Next Generation of RL, NGRL

skills/multi-task transfer
agent should share skills and understanding across scenario
from Montezuma's revenge to continuous action space sparse reward robotics, with WM&CL

Exponential NxG RL (next next?)

Algorithm

O(1), O(n), O(nlogn), O(n^2).. .

Neural Netwc

train

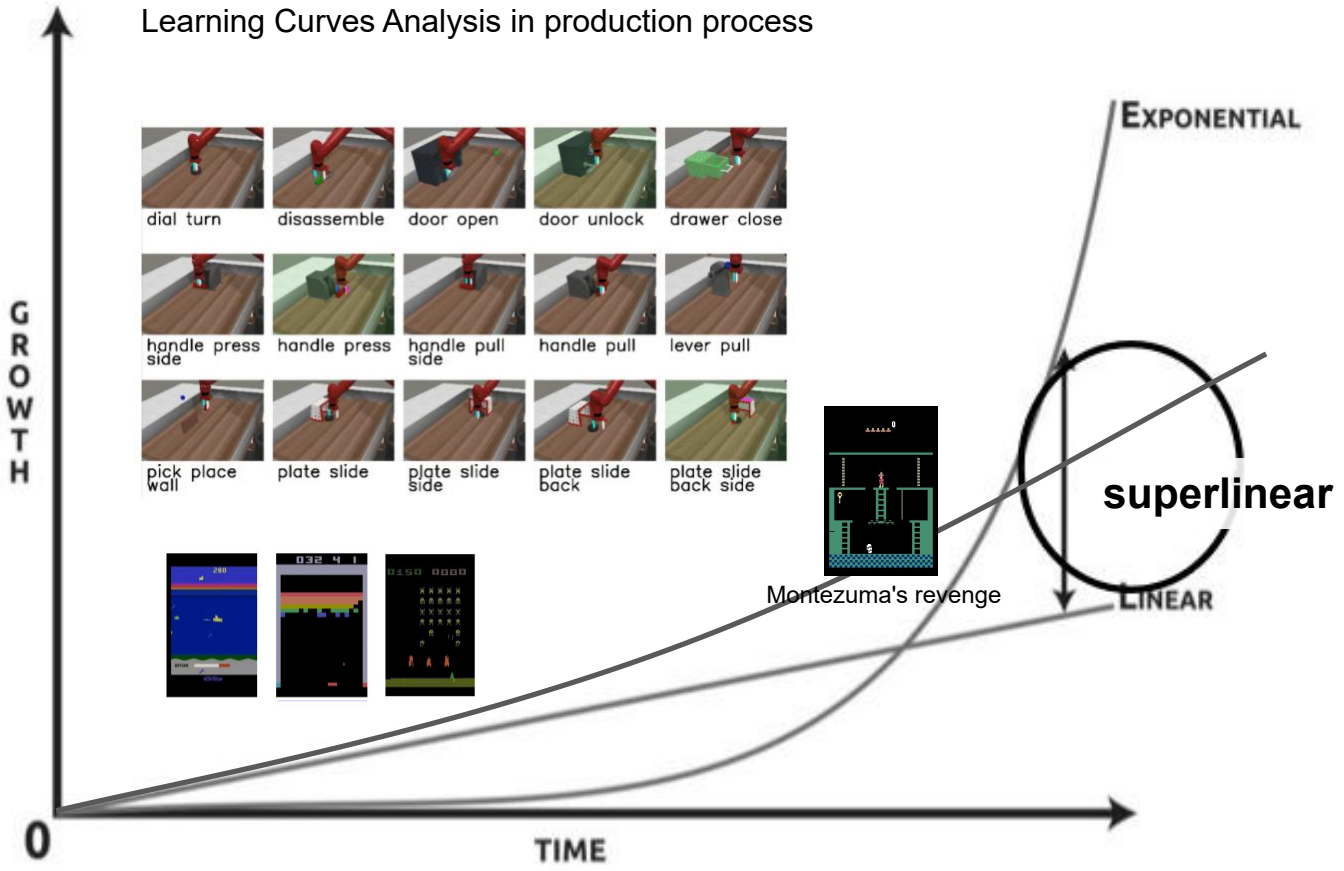
few-

one-shot learning

zero-shot learning

.. .

$$\left(e^x\right)' = e^x$$



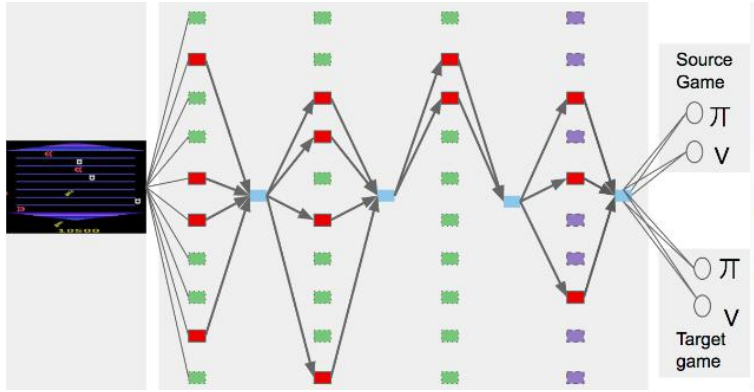
fundamentally, experiences/skills should accelerate the learning process across the tasks/environment

with more experiences, the agent should learn the task more fast.

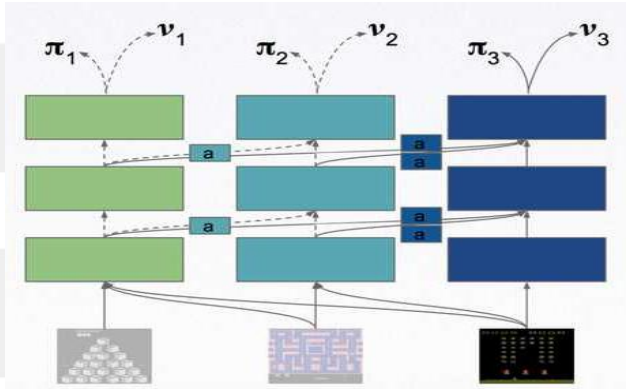
If it's proportional to the experiences,

by definition, it should be an exponential learning curve.

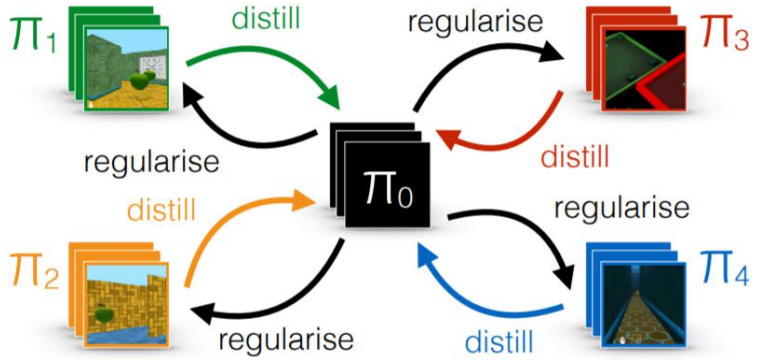
Multi-task RL Exist now



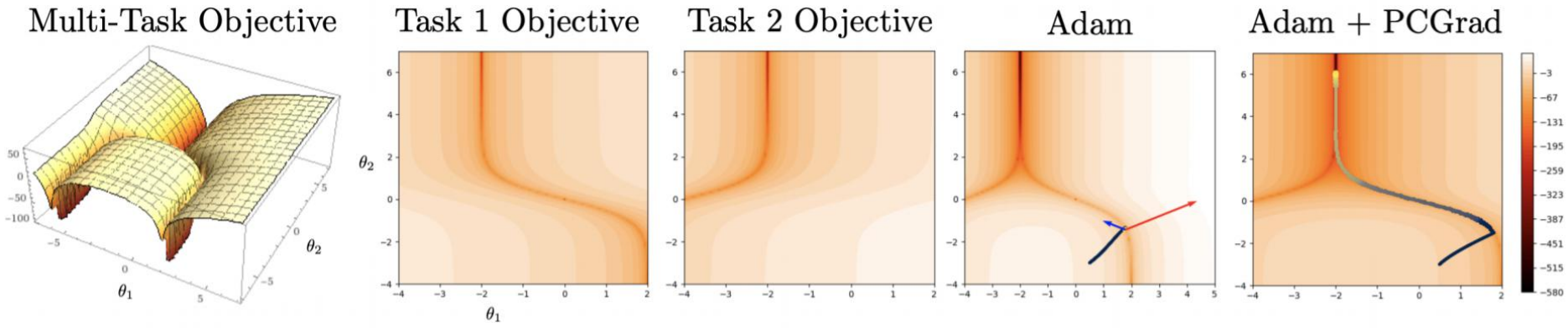
PathNet



Progressive Neural Networks

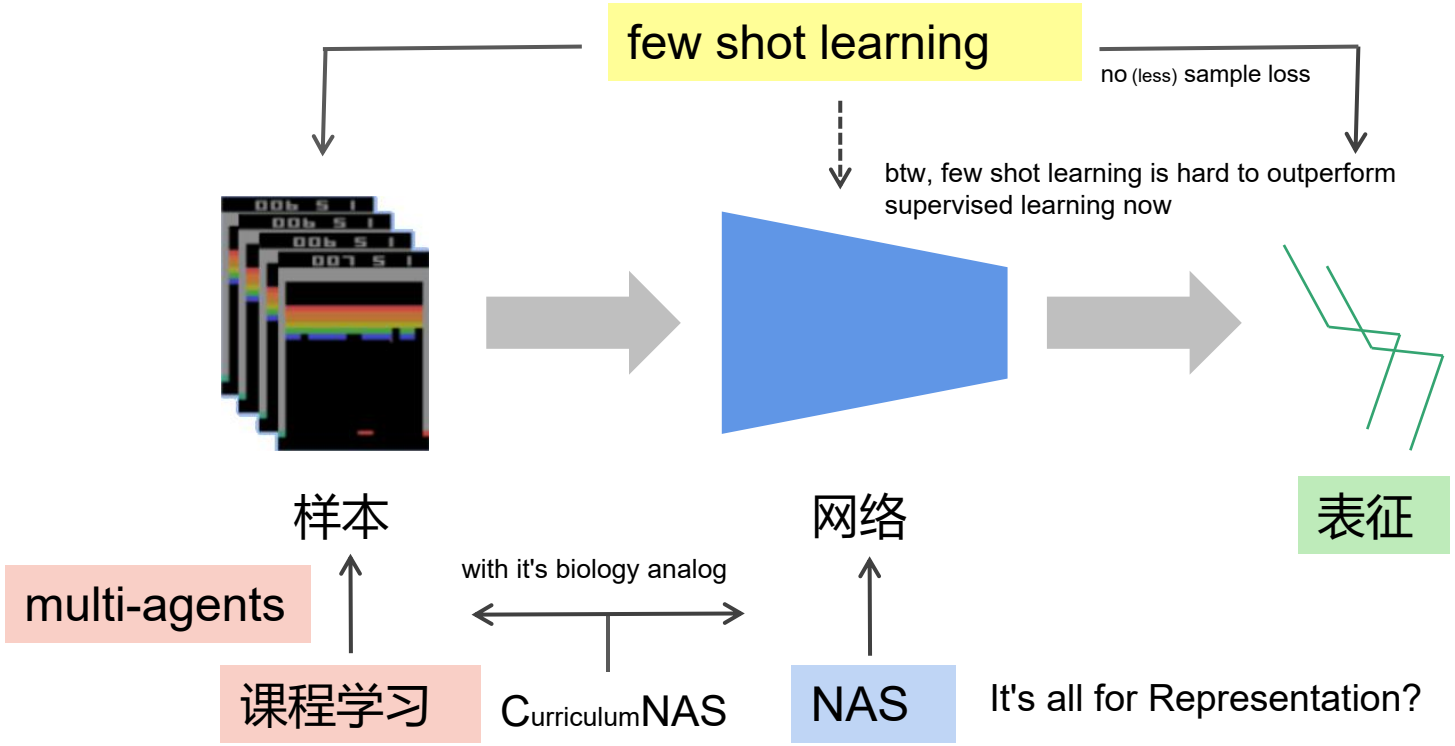


Distral: Distill & transfer learning



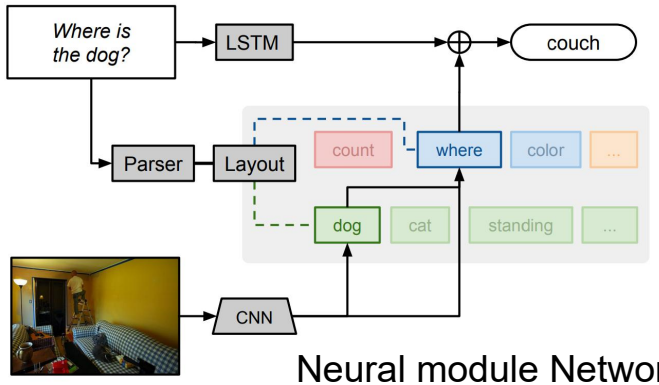
Project Conflicting Gradients
PopArt Normalization

Let’s go back to the Neural Networks

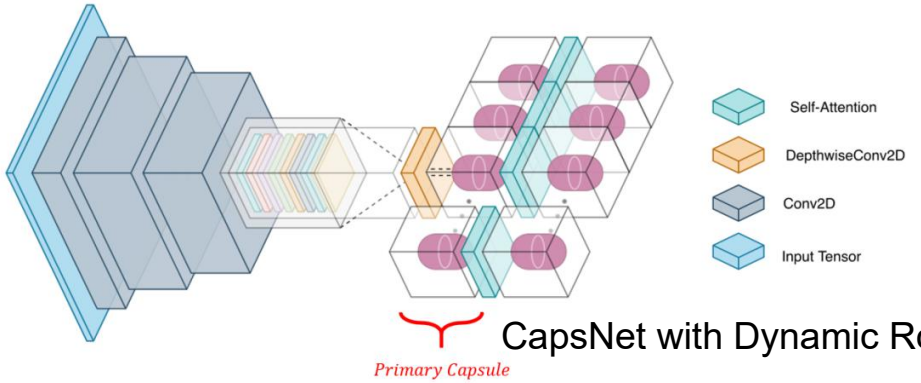


Breaking the Curse of Space Explosion, 2020, Tencent, arXiv: 2007.07197

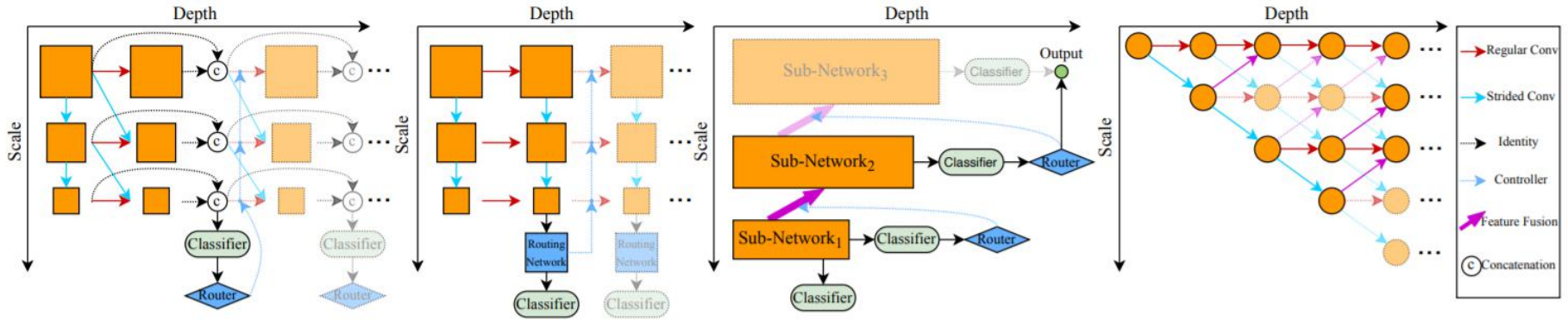
Dynamic Networks available now



Neural module Network for VQA

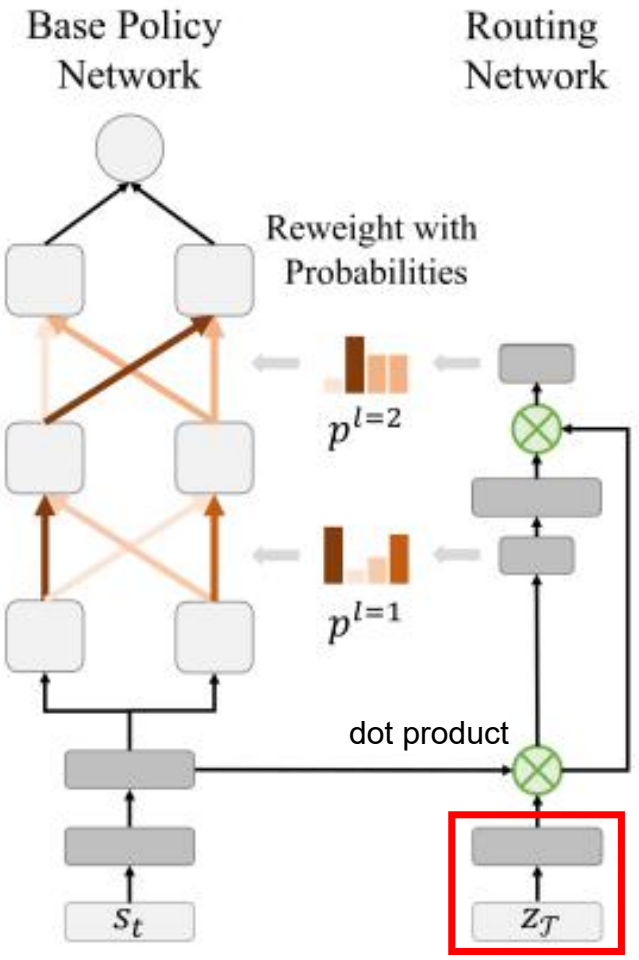


CapsNet with Dynamic Routing



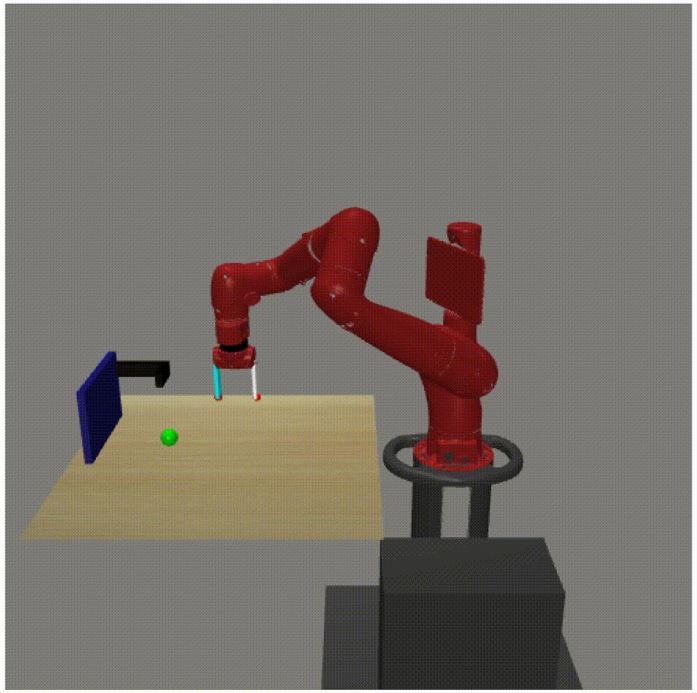
Dynamic Neural Networks

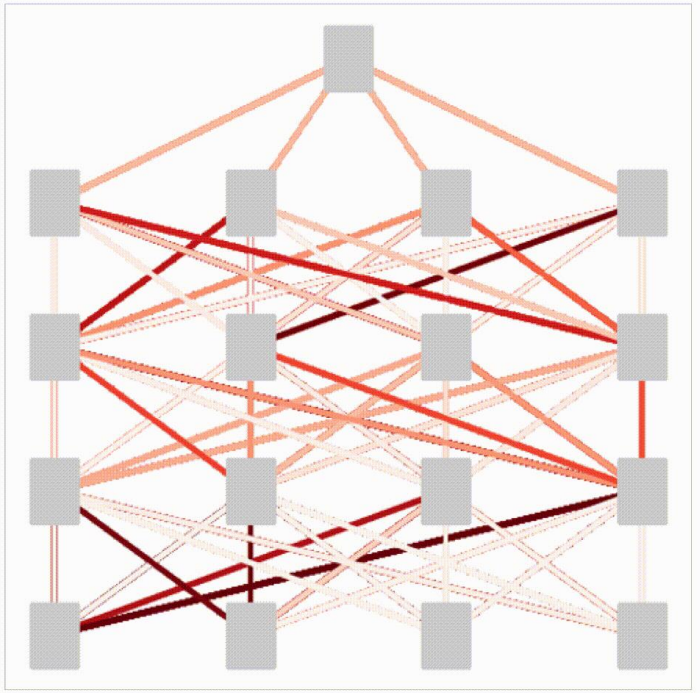
Mostly, design for visual task

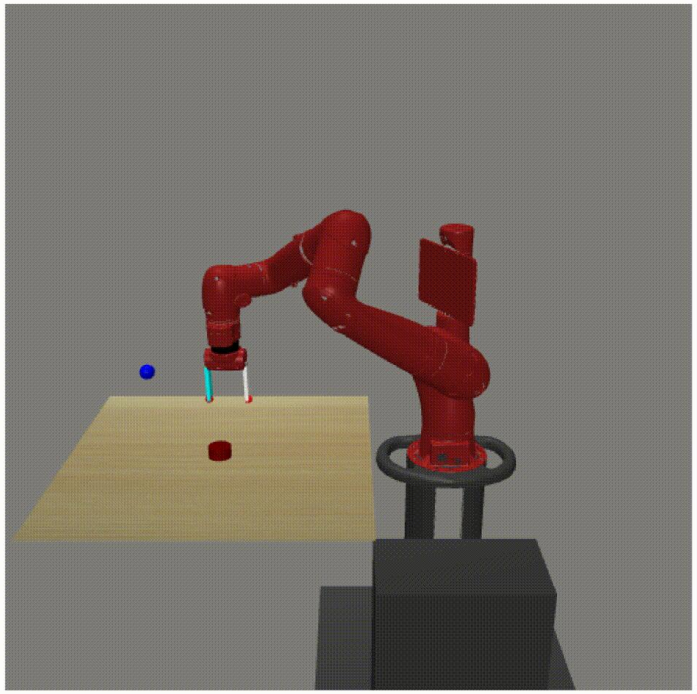


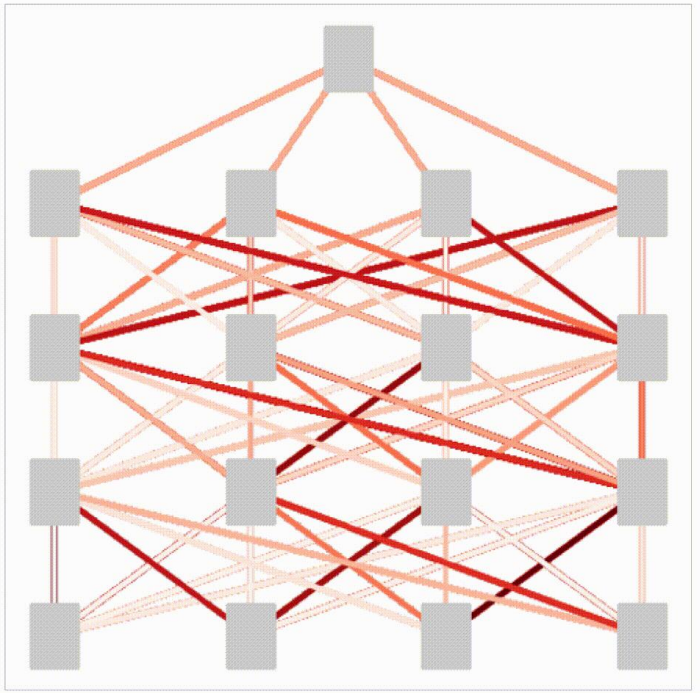
Multi-Task Reinforcement Learning with Soft Modularization,

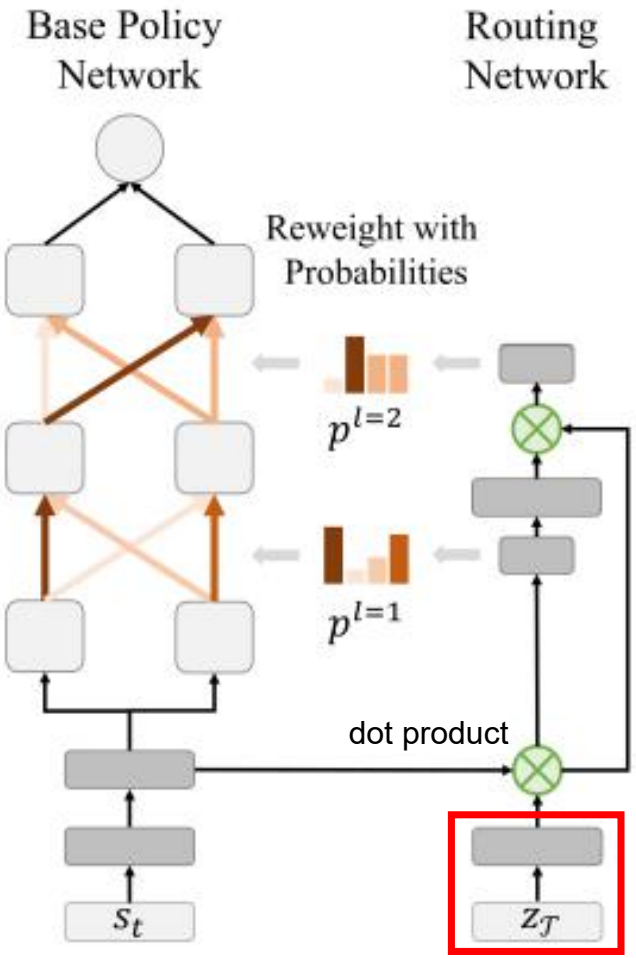
NIPS2020, arxiv2003.13661











one-hot task embedding

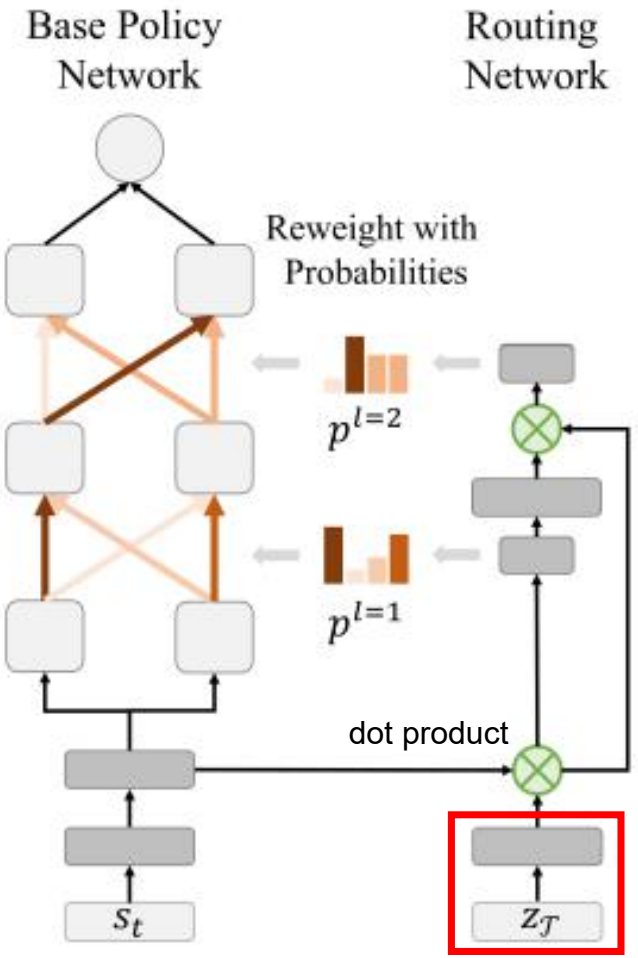
Embedding
- together with Time-series?
- better Task Embedding? including Hierarchical Task Structure?
- Processes of the task? relationship between the Options?

Gating
- a new Attention mechanism on Network Architecture?
- dynamic NAS?
- Sparsity Gating (to save computing power)?

Modularity
- What do the Soft Modules Stand for?
- How they get functionalilty across tasks
- How to generalize this mechanism to other part of the agent?
  - like Value networks, World Models embedding, etc

a Paradigm for Continue/Lifelong Learning?

topics on ICRA2022



one-hot task embedding

Embedding

together with WorldModel Time-series?

better Task Embedding? including Hierarchical Task Structure?

Processes of the task? relationship between the Options?

Gating

a new Attention mechanism on Network Architecture?

dynamic NAS?

Sparsity Gating (to save computing power)?

Modularity

What do the Soft Modules Stand for?

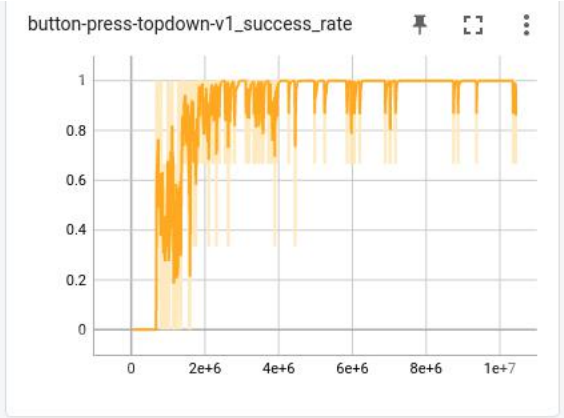
How they get functionalilty across tasks

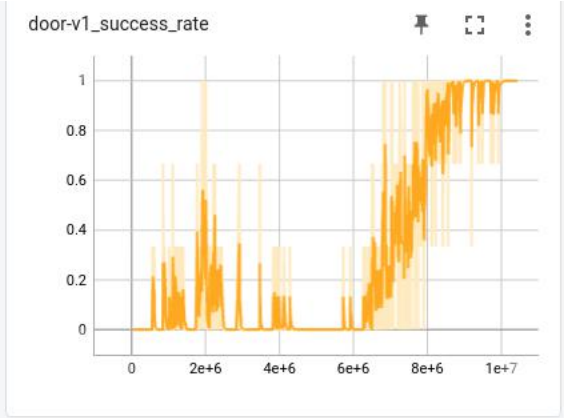
How to generalize this mechanism to other part of the agent?

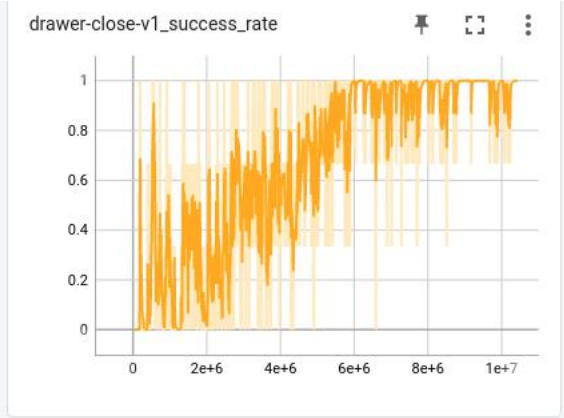
like Value networks, World Models embedding, etc

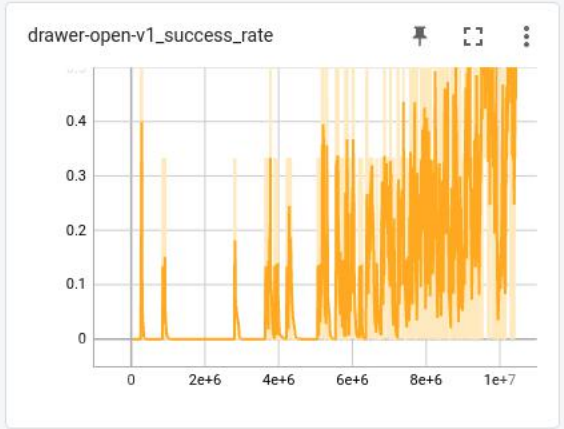
a Paradigm for Continue/Lifelong Learning?

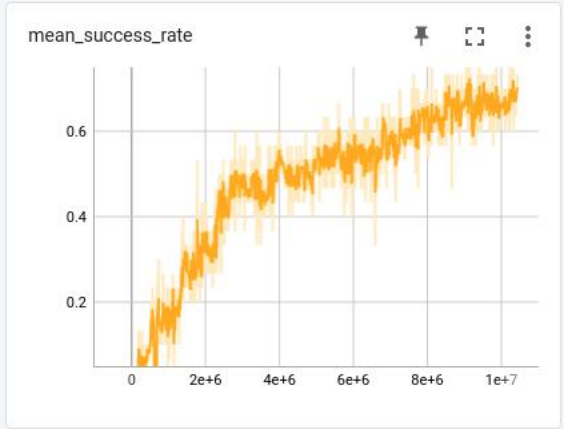
Primary result on Soft Modularization

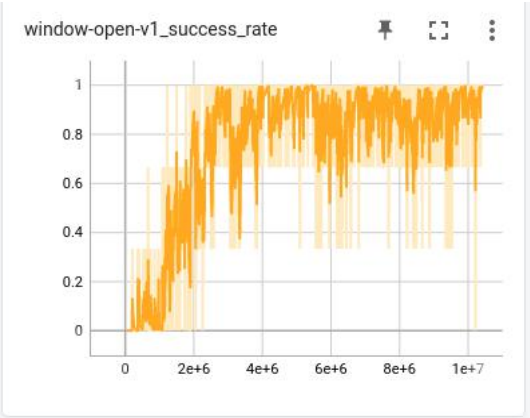


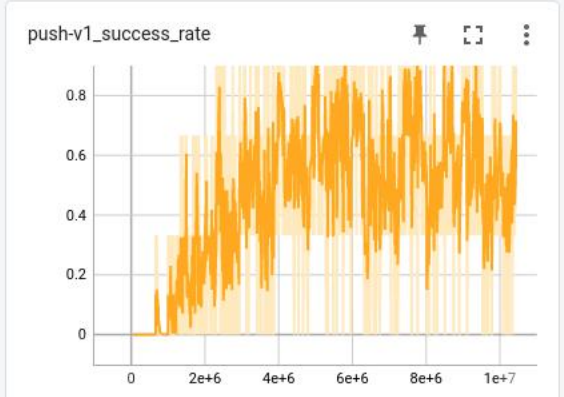


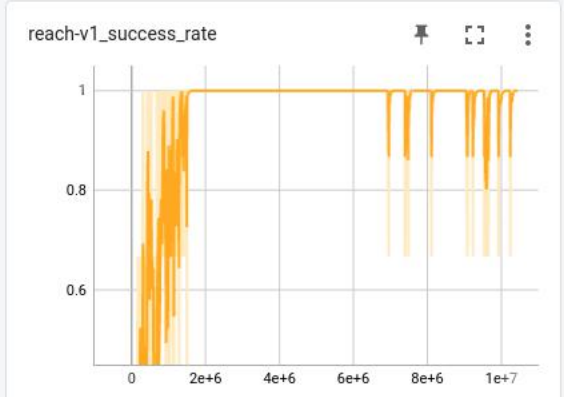


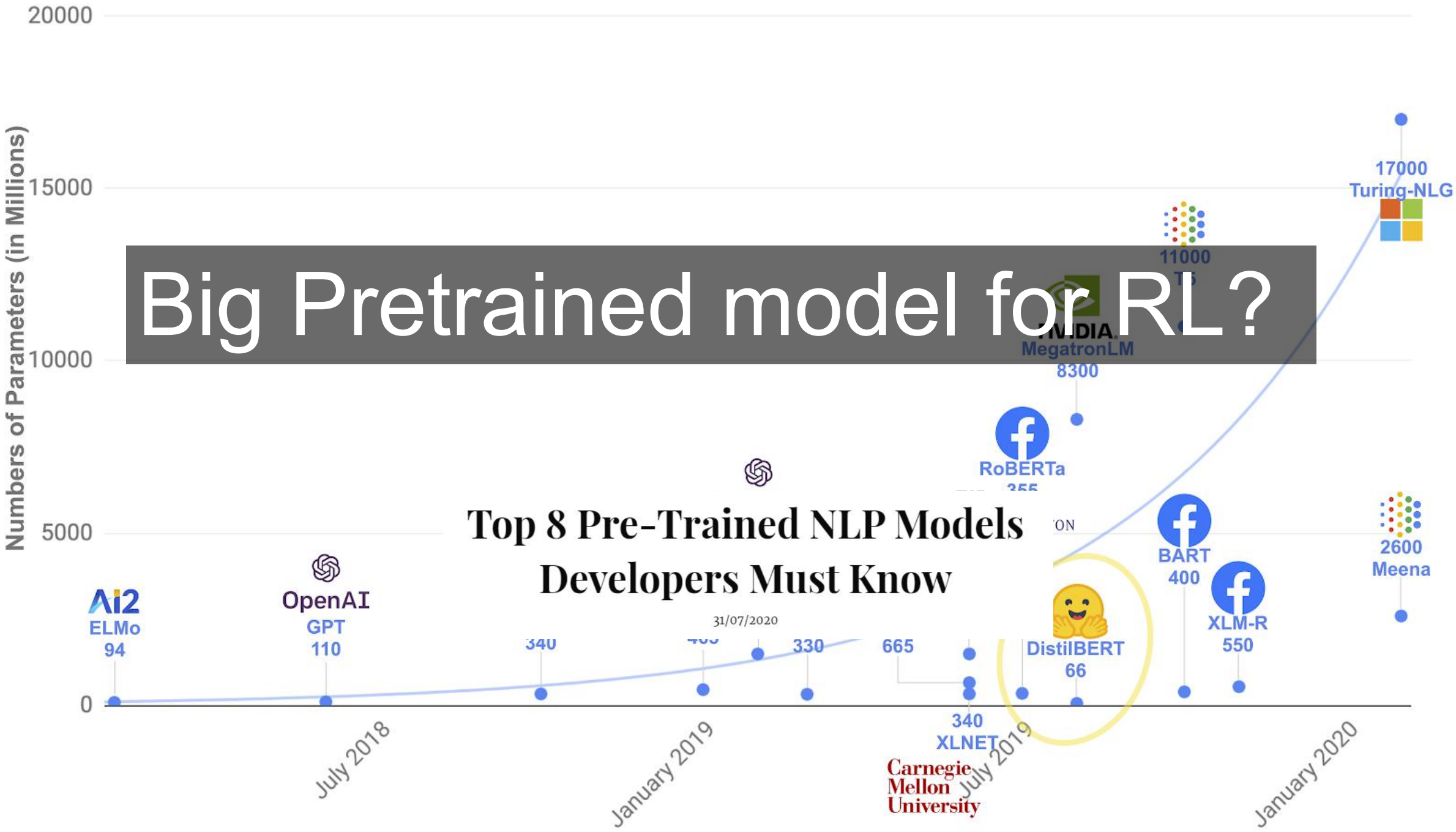











Big Pretrained model for RL?
Top 8 Pre-Trained NLP Models Developers Must Know
31/07/2020
Numbers of Parameters (in Millions)
20000
15000
10000
5000
0
July 2018
January 2019
July 2019
January 2020
Ai2
ELMo
94
OpenAI
GPT
110
340
330
665
340
XLNET
Carnegie Mellon University
RoBERTa
MegatronLM
8300
11000
17000
Turing-NLG
DistilBERT
66
BART
400
XLM-R
550
2600
Meena

Thanks!