UNIVERSITY OF LATVIA
FACULTY OF COMPUTING

Deep Learning for Systematic Generalization

Guntis V. Strazds

guntis_vilnis.strazds@lu.lv

December 22, 2021

Supervisor: Prof. Guntis Bārzdiņš, Dr.sc.comp

Systematic Generalization, what does this mean?

- Algorithms (fully systematic)
- Classification
  - Categorization (can be formal or "fuzzy": prototypes, etc)
  - Inheritance, is-a relationships
- Analogies
  - X (in some context) can be viewed as ...
  - X (in some sense) behaves similarly to ...

Proving a theorem requires algorithmic and categorical reasoning. Theorizing (inventing/discovering a new theorem or theory) often involves analogical thinking. Natural language makes extensive use of all of the above.

Introduction

Generalization and Systematicity

Santoro et al. 2021 - **Symbolic Behavior in Artificial Intelligence** [14]

d'Avila Garcez & Lamb 2020 - **Neurosymbolic AI: The 3rd Wave** [5]

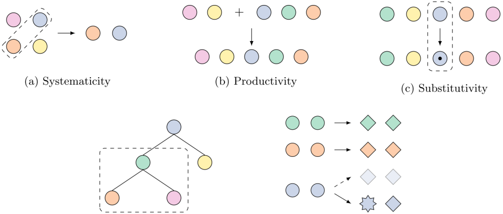
Goyal & Bengio 2020 - **Inductive Biases for Deep Learning of Higher-Level Cognition** [6]

Hupkes et al. 2020 - **Compositionality Decomposed: How do Neural Networks Generalise?** [10]

Kirk et al. 2021 - **A Survey of Generalisation in Deep Reinforcement Learning** [12]

Introduction

Types of Generalization

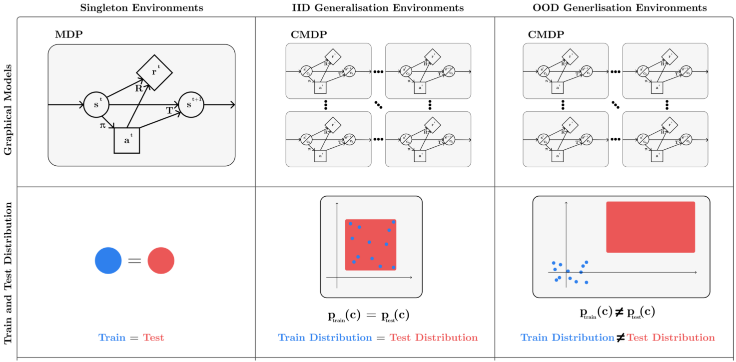


(a) Systematicity (b) Productivity (c) Substitutivity (d) Localism (e) Overgeneralisation

Source: Hupkes et al. 2020 - Compositionality Decomposed: How do Neural Networks Generalise? [10]

	Definition
(a) Systematicity	Recombine constituents that have not been seen together during training
(b) Productivity	Test sequences longer than ones seen during training
(c) Substitutivity	Meaning unchanged if a constituent is replaced with something equivalent
(d) Localism	The meaning of local parts are unchanged by the global context
(e) Overgeneralization	Can handle exceptions to rules and patterns?

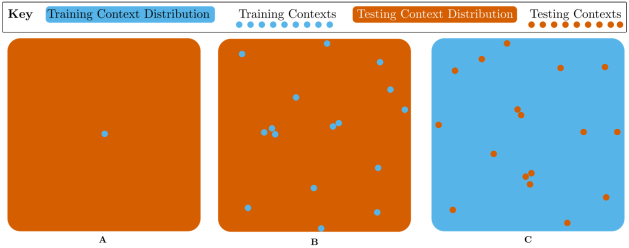
Definitions from: `https://evjang.com/2021/12/17/lang-generalization.html`

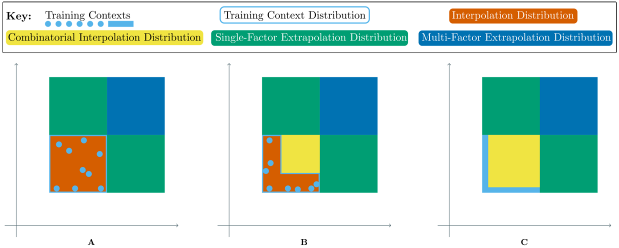


Kirk et al. 2021 - A Survey of Generalisation in Deep Reinforcement Learning [12]

Introduction

Generalization in Deep RL





Kirk et al. 2021 - A Survey of Generalisation in Deep Reinforcement Learning [12]

Environments for Experiments

- MiniGrid
  - `https://github.com/maximecb/gym-minigrid`
- TextWorld
  - `https://github.com/microsoft/TextWorld`
  - Microsoft Research blog about TextWorld
- MiniHack
  - `https://github.com/facebookresearch/minihack`
  - Facebook AI Research blog about MiniHack

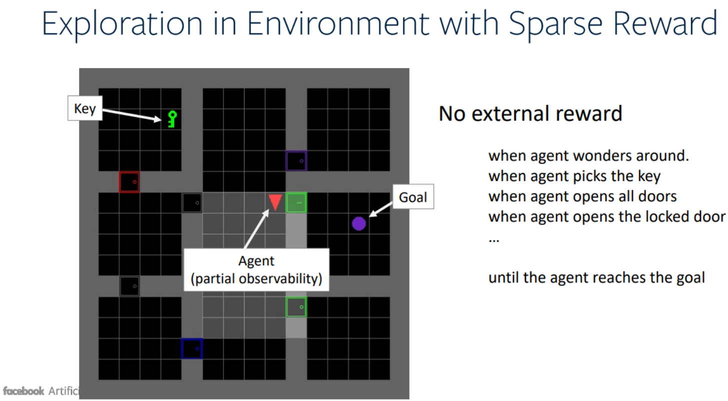


Illustration from `https://https://yuandong-tian.com/ucl_dark_talk_2021.pdf`

And more complicated situations...

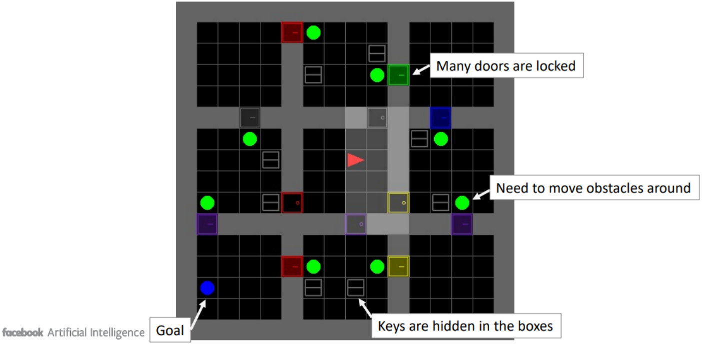
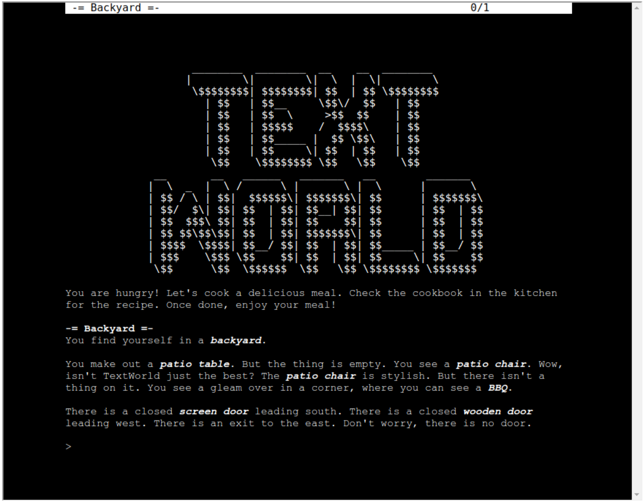


Figure reproduced from `https://https://yuandong-tian.com/ucl_dark_talk_2021.pdf`

Environments

TextWorld – A Platform for Text Adventure Games



https://www.microsoft.com/en-us/research/project/textworld/try-it/

Environments
TextWorld – Game Generation

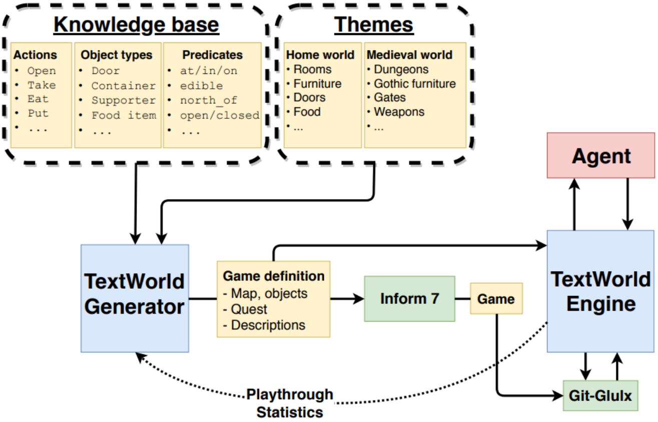


Fig.3 from Côté et al. 2019 - **TextWorld: A Learning Environment for Text-Based Games** [3]

(click here for list of available challenges)

Environments

MiniHack

see an animated example

Learning Systematic Action Policies

How can we build and train models that learn to act in ways that **generalize** to previously unseen environments or situations?

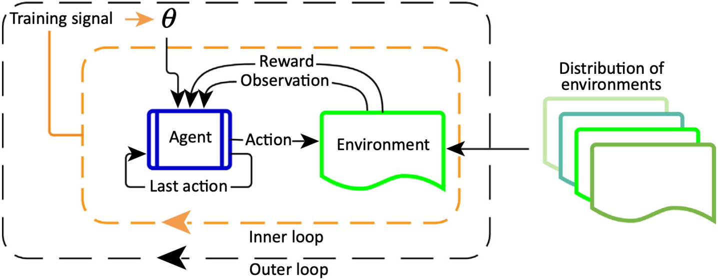
And that display **compositional systematicity**, e.g. by being able to reuse sub-skills in new combinations when appropriate?

Learning Systematic Action Policies

(Reinforcement-/Imitation-/Meta-/Continual-/Curriculum-/...etc)

- Meta-Learning
- Imitation Learning and Offline RL
- Continual / Lifelong Learning
- Curriculum Learning (curiosity = auto-curriculum)
- Hindsight Experience Replay
- Learning environment dynamics; Model-based RL, latent space planning
- State Representation (consolidating, long trajectories)
  - Explicit long-term memory, KGs
- Hierarchical RL
- Hybrid/neuro-symbolic, neuro-algorithmic

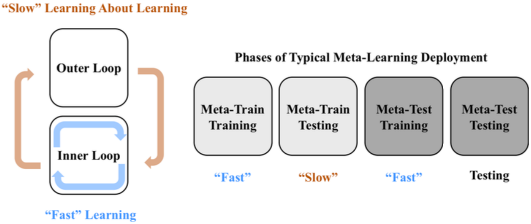
Learning Systematic Action Policies
Meta-Learning

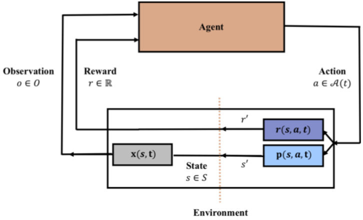


Copied from Fig.1 of Botvinik et al. 2019 - **Reinforcement Learning, Fast and Slow** [1]

Learning Systematic Action Policies

Meta-Learning: Meta-Training, Meta-Testing





Khetarpal et al. 2020 - **Towards Continual Reinforcement Learning A Review and Perspectives** [11]

Learning Systematic Action Policies

Continual RL and Transfer-Learning

Setting	Multiple Domains Of Deployment	Multiple Required Skills	Universal Master Policy	Non-stationary Evolution
Domain Adaptation	✓	X	X	X
Transfer Learning	✓	✓	X	✓
Meta-Training and Meta-Testing	✓	✓	X	X
Multi-task Learning	✓	✓	✓	X
Continual (Lifelong) Learning	✓	✓	✓	✓

Khetarpal et al. 2020 - **Towards Continual Reinforcement Learning A Review and Perspectives** [11]

Learning Systematic Action Policies

Continual Reinforcement Learning

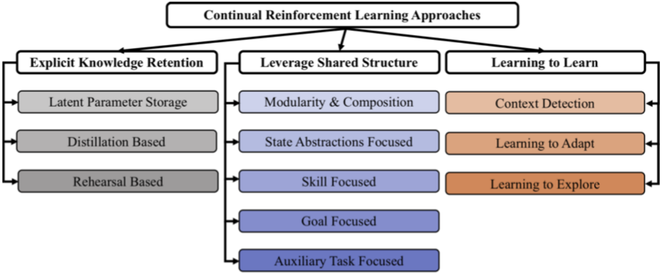
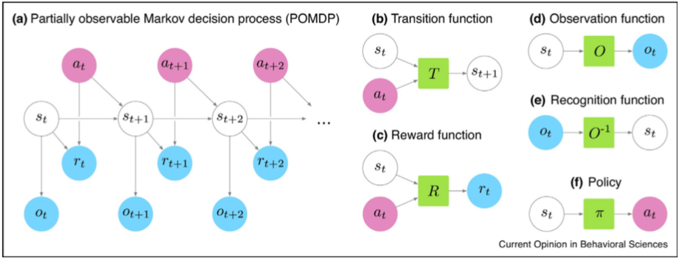


Figure 5: **Taxonomy of Continual RL Approaches:** A diagram illustrating different clusters of approaches for continual RL, highlighting prominent threads of research within each family.

Khetarpal et al. 2020 - **Towards Continual Reinforcement Learning A Review and Perspectives** [11]

Learning Systematic Action Policies

Model-based RL and Planning



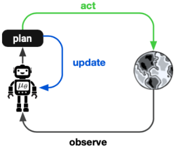
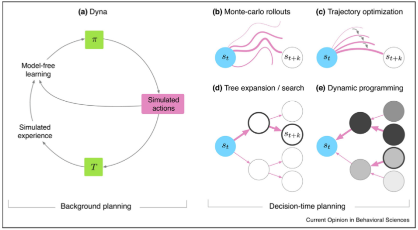


Figure 1: Model-based approximate policy iteration. The agent **updates** its policy using targets computed via planning and optionally **acts** via planning during training, at test time, or both.



From Hamrick 2019 - **Analogues of mental simulation and imagination in deep learning**[7]
And: Hamrick et al. 2020 - **On the role of planning in model-based deep reinforcement learning** [9]

Learning Systematic Action Policies

Simulation-based inference

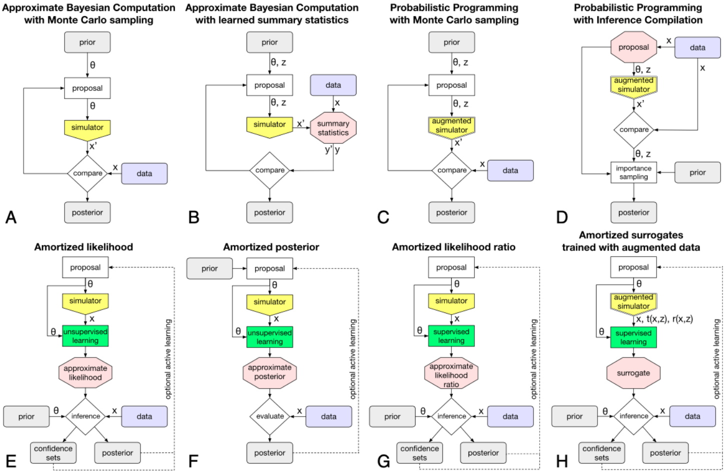


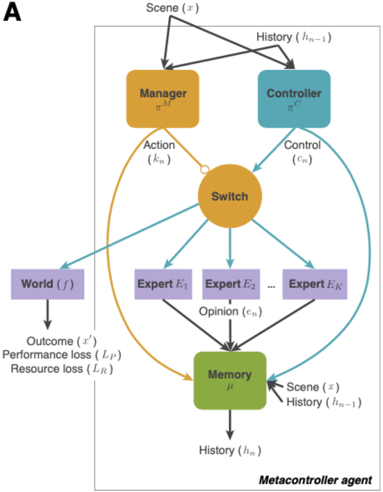
Fig. 1. *(A–H)* Overview of different approaches to simulation-based inference.

Cranmer et al. 2020 - **The frontier of simulation-based inference** [4]
(See also: Mohamed et al. 2020 - **Monte Carlo Gradient Estimation in Machine Learning** [13])

Learning Systematic Action Policies

Modular and Hierarchical RL, Mixture of Experts

1. Manager = Metacontroller
2. Controller = Action selection
3. Experts = Simulator modules
4. Switch: routes actions to a simulator or to the World



Hamrick et al. 2017 - **Metacontrol for Adaptive Imagination-Based Optimization** [8]

Veličković & Blundell - 2021 - **Neural algorithmic reasoning** [15]

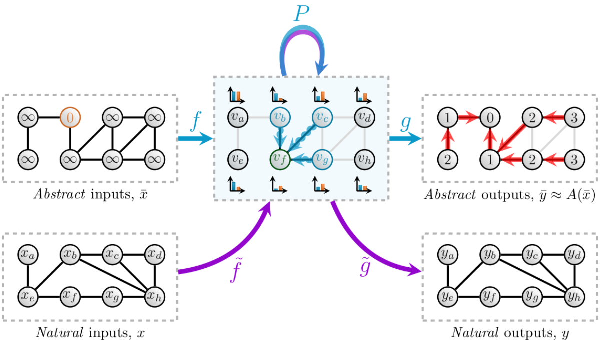


Figure 7: The proposed algorithmic reasoning blueprint. First, an algorithmic reasoner is trained in the encode-process-decode fashion, learning a function $g(P(f(\bar{x}))) \approx A(\bar{x})$, for a target combinatorial algorithm A; in this case, A is breadth-first search. Once trained, the processor network P is frozen and stitched into a pipeline over natural inputs—with new encoder and decoder $\tilde{f}$ and $\tilde{g}$. This provides an end-to-end differentiable function that has no explicit information loss, while retaining alignment with BFS.

Reproduced from Cappart et al. 2021 [2]

Tools and Frameworks

What I'm currently doing: updating my workbench

- Huggingface Transformers, Datasets
- SaLinA: Sequential Learning of Agents
- FAIR xFormers

Ideas & Research Goals

"This might work..."

Learning Systematic Action Policies

Note on Mega-scaling vs. Inductive biases

Sutton's "bitter lesson from 70 years of AI research"[1]
Given exponentially increasing computing resources, general purpose learning and search methods end up, over a time span only slightly longer than a typical research project, outperforming knowledge-intensive, hand-crafted approaches.[2]

But the scale of many current SoA models is now beyond the reach of most academic researchers. So what can we do?

[1] http://www.incompleteideas.net/IncIdeas/BitterLesson.html

[2] See this blog for counterpoint re successful algorithmic methods based on human understanding

Ideas & Research Goals

Using ideas from Veličković & Blundell - 2021 - **Neural algorithmic reasoning** [15, 16]

- Modular neuro-algorithmic software
- Mimicking the general structure of TW Oracle
- Demonstrate:
  1. Systematic and productive generalization (by reliably solving TextWorld tasks)
  2. Transfer-adaptation: (by learning to interpret 'noisy' raw textual observations by training outer model with neuro-algorithmic core trained using 'clean' ground-truth or semantic observations)
  3. Generalization from TextWorld to MiniGrid (by adapting or retraining a subset of sub-modules)

References I

[1] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, and D. Hassabis. Reinforcement Learning, Fast and Slow. *Trends Cogn. Sci.*, xx, 2019.

[2] Q. Cappart, D. Chételat, E. Khalil, A. Lodi, C. Morris, and P. Veličković. Combinatorial optimization and reasoning with graph neural networks. feb 2021.

[3] M. A. Côté, Á. Kádár, X. Yuan, B. Kybartas, T. Barnes, E. Fine, J. Moore, M. Hausknecht, L. El Asri, M. Adada, W. Tay, and A. Trischler. TextWorld: A Learning Environment for Text-Based Games. In *Commun. Comput. Inf. Sci.*, volume 1017, pages 41–75, jun 2019.

References II

[4] K. Cranmer, J. Brehmer, and G. Louppe. The frontier of simulation-based inference. *Proc. Natl. Acad. Sci. U. S. A.*, 117(48):30055–30062, dec 2020.

[5] A. d'Avila Garcez and L. C. Lamb. Neurosymbolic AI: The 3rd Wave. dec 2020.

[6] A. Goyal and Y. Bengio. Inductive Biases for Deep Learning of Higher-Level Cognition. *arXiv*, nov 2020.

[7] J. B. Hamrick. Analogues of mental simulation and imagination in deep learning, 2019.

[8] J. B. Hamrick, A. J. Ballard, R. Pascanu, O. Vinyals, N. Heess, and P. W. Battaglia. Metacontrol for Adaptive Imagination-Based Optimization. In *ICLR*, may 2017.

UNIVERSITY OF LATVIA

References III

[9] J. B. Hamrick, A. L. Friesen, F. Behbahani, A. Guez, F. Viola, S. Witherspoon, T. Anthony, L. Buesing, P. Veličković, and T. Weber. On the role of planning in model-based deep reinforcement learning. *arXiv*, nov 2020.

[10] D. Hupkes, V. Dankers, M. Mul, and E. Bruni. Compositionality Decomposed: How do Neural Networks Generalise? *J. Artif. Intell. Res.*, 67:757–795, 2020.

[11] K. Khetarpal, M. Riemer, I. Rish, and D. Precup. Towards Continual Reinforcement Learning: A Review and Perspectives. dec 2020.

[12] R. Kirk, A. Zhang, E. Grefenstette, and T. Rocktäschel. A Survey of Generalisation in Deep Reinforcement Learning. nov 2021.

UNIVERSITY OF LATVIA
FACULTY OF COMPUTING

References IV

[13] S. Mohamed, M. Rosca, M. Figurnov, and A. Mnih. Monte Carlo Gradient Estimation in Machine Learning. *J. Mach. Learn. Res.*, 21:1–63, jun 2020.

[14] A. Santoro, A. Lampinen, K. Mathewson, T. Lillicrap, and D. Raposo. Symbolic Behaviour in Artificial Intelligence. feb 2021.

[15] P. Veličković and C. Blundell. Neural algorithmic reasoning. *Patterns*, 2(7), may 2021.

[16] P. Veličković, M. Bošnjak, T. Kipf, A. Lerchner, R. Hadsell, R. Pascanu, and C. Blundell. Reasoning-Modulated Representations. jul 2021.

Thank You.

(to be continued...)

guntis_vilnis.strazds@lu.lv