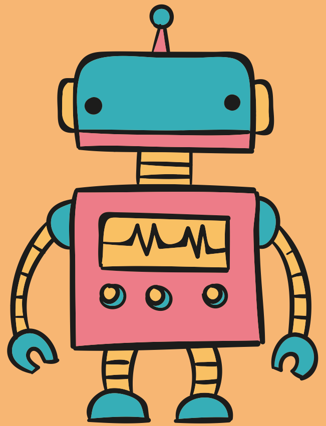
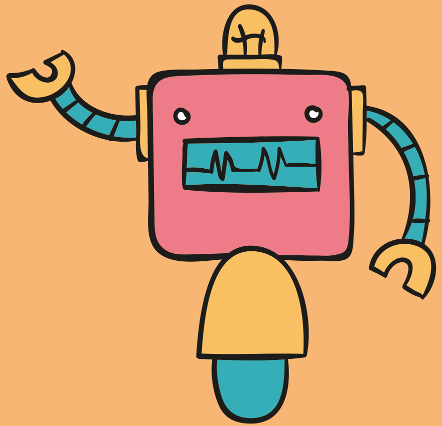
Towards Vygotskian Autotelic Agents

Learning Skills with Goals, Language and Intrinsically Motivated Reinforcement Learning



Cédric Colas

Advisors: Pierre-Yves Oudeyer and Olivier Sigaud



Content

Piagetian
Autotelic Learning

Vygotskian
Autotelic Learning

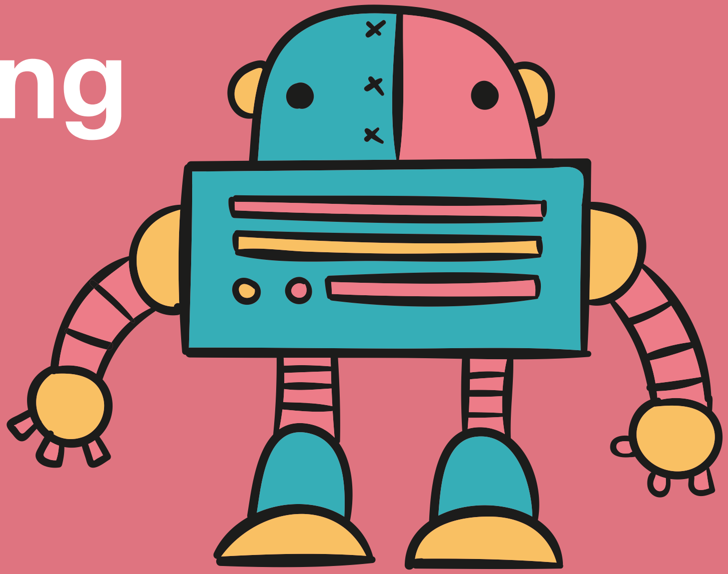
Study #1
IMAGINE

Study #2
DECSTR

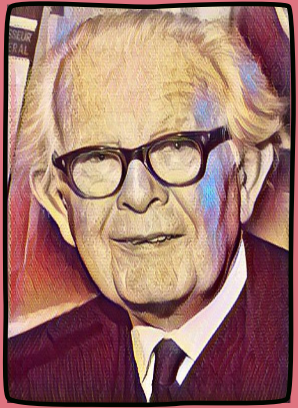
Perspectives

Piagetian Autotelic Learning

Concepts and related work



Intrinsically Motivated Learners



Jean Piaget (1896-1980)

Credit: Francis Vachon

Intrinsic motivations

Defined by psychologists
(Berlyne, 1950/1966; Czikszentmihalyi, 1990; Ryan & Deci, 2000; Kidd, 2012).

Implemented by reinforcement learning and developmental robotics researchers
(Schmidhuber, 1991; Oudeyer & Kaplan, 2004, 2007).

Scaled with deep RL (knowledge-based)
(Bellemare, 2016; Pathak, 2018; Burda, 2019).

Intrinsically Motivated Goal-Directed Learners

Open-ended repertoire of skills

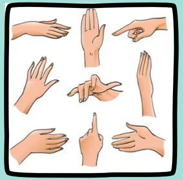
Goals

"A goal is a cognitive representation of a future object that the organism is committed to approach or avoid."
(Elliott & Fryer, 2008)

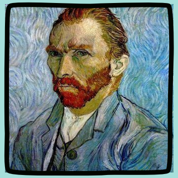
$$g = (z_g,\ R_g)$$

Diversity of goal representations (modalities, abstraction)

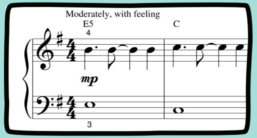
Proprioceptive



Visual



Auditory



Linguistic

Do good research

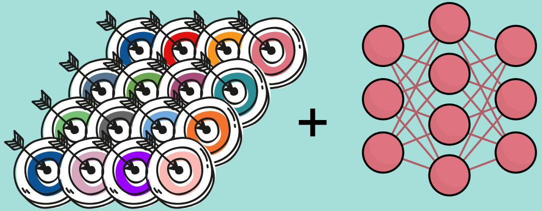
Autotelic Learners

Autotelic Learning

Autotelic agents are intrinsically motivated to learn to represent, generate, pursue and master their own goals.

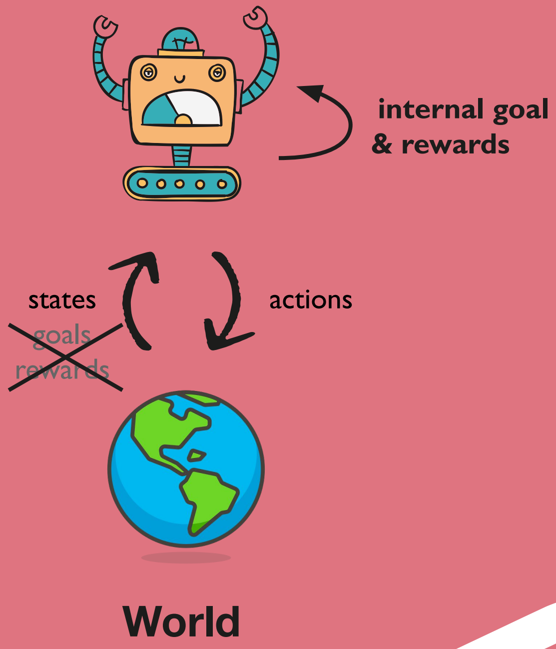
"Autotelic" comes from the Greek *auto* (self) and *telos* (end, goal) (Steels, 2004).

Repertoire of Skills

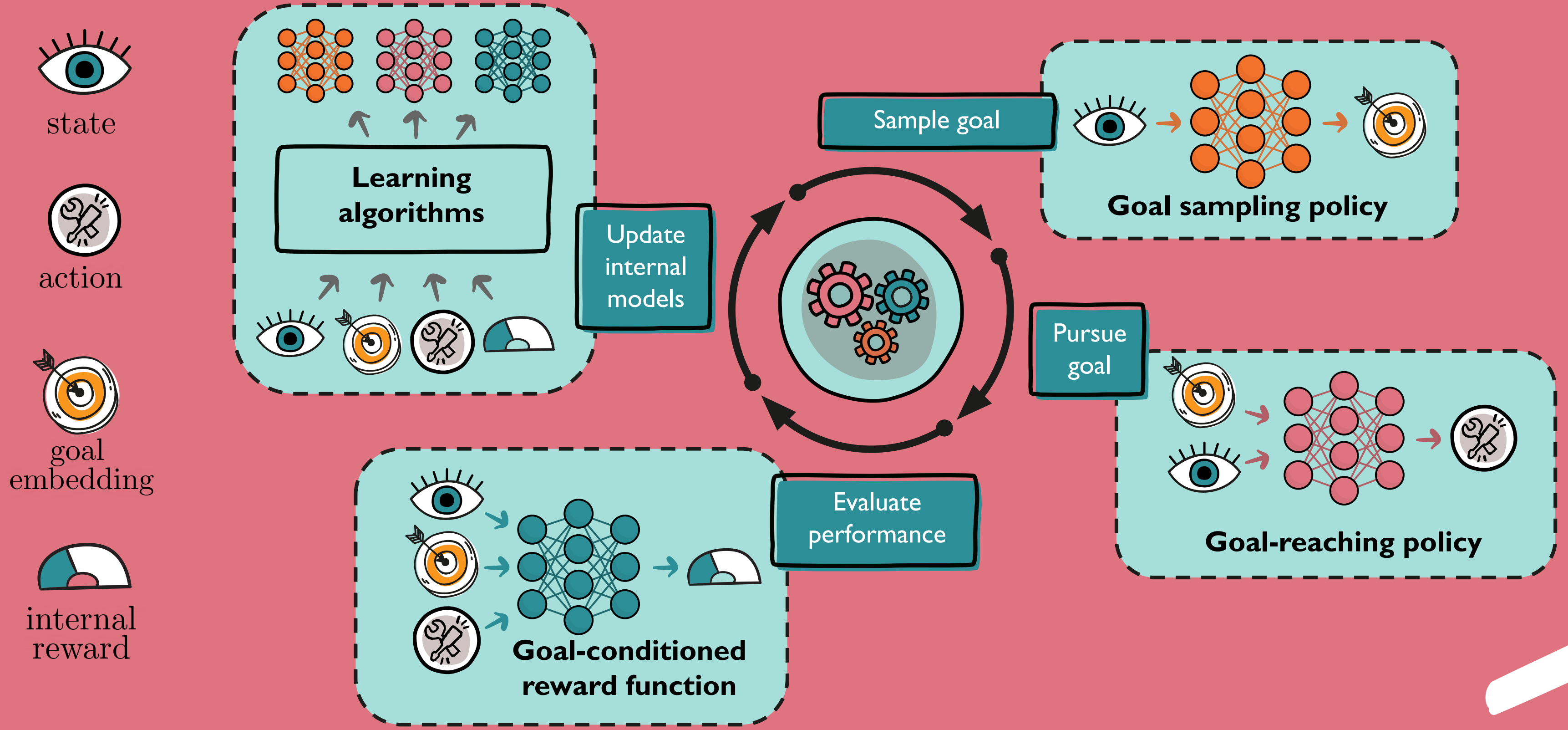


set of goal representations + goal-directed behaviors

Autotelic agent

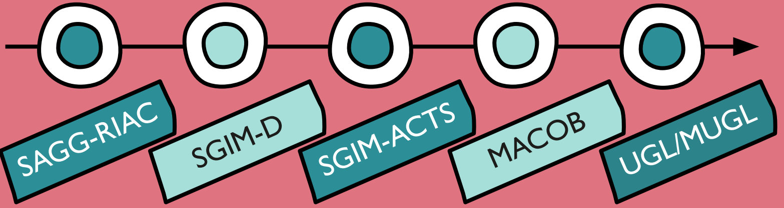


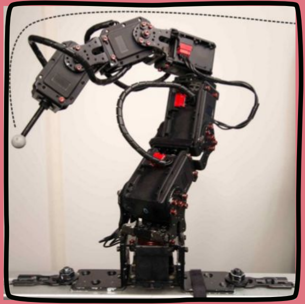
Autotelic Learning Loop

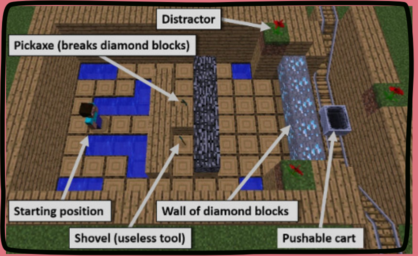


Autotelic Learning with IMGEPs

Intrinsically Motivated Goal Exploration Processes (IMGEPs) implement autotelic learning with competence-based intrinsic motivations.







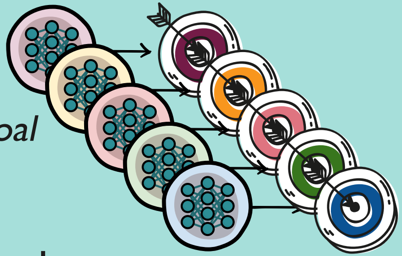
Algorithm families	**Learning to represent goals**	**Learning to sample goals**	**Learning to evaluate goal-reaching**	**References**
Existing IMGEPs	Hand-coded representations ✘	Learning progress	Hand-coded reward functions ✘	Baranes, 2010; Nguyen, 2011/2012 Moulin-Frier, 2014; Forestier, 2016.

SAGG-RIAC (Baranes, 2010); SGIM-D/SGIM-ACTS (Nguyen 2011/2012), MACOB (Forestier, 2016), Benureau & Oudeyer, 2016 UGL (Péré, 2018), MUGL (Laversanne-Finot, 2018)

POP-IMGEPs and Reinforcement Learning

POP-IMGEPs

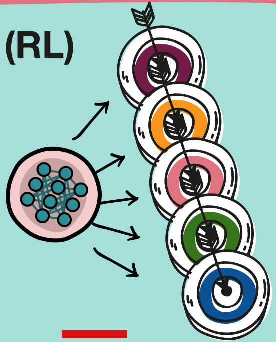
one policy → one goal



+	−
Simple Good exploration Fast for coarse solutions No diff. requirements Parallelizable	Poor finetuning Poor generalization Low sample efficiency Memory (archive)

Reinforcement Learning (RL)

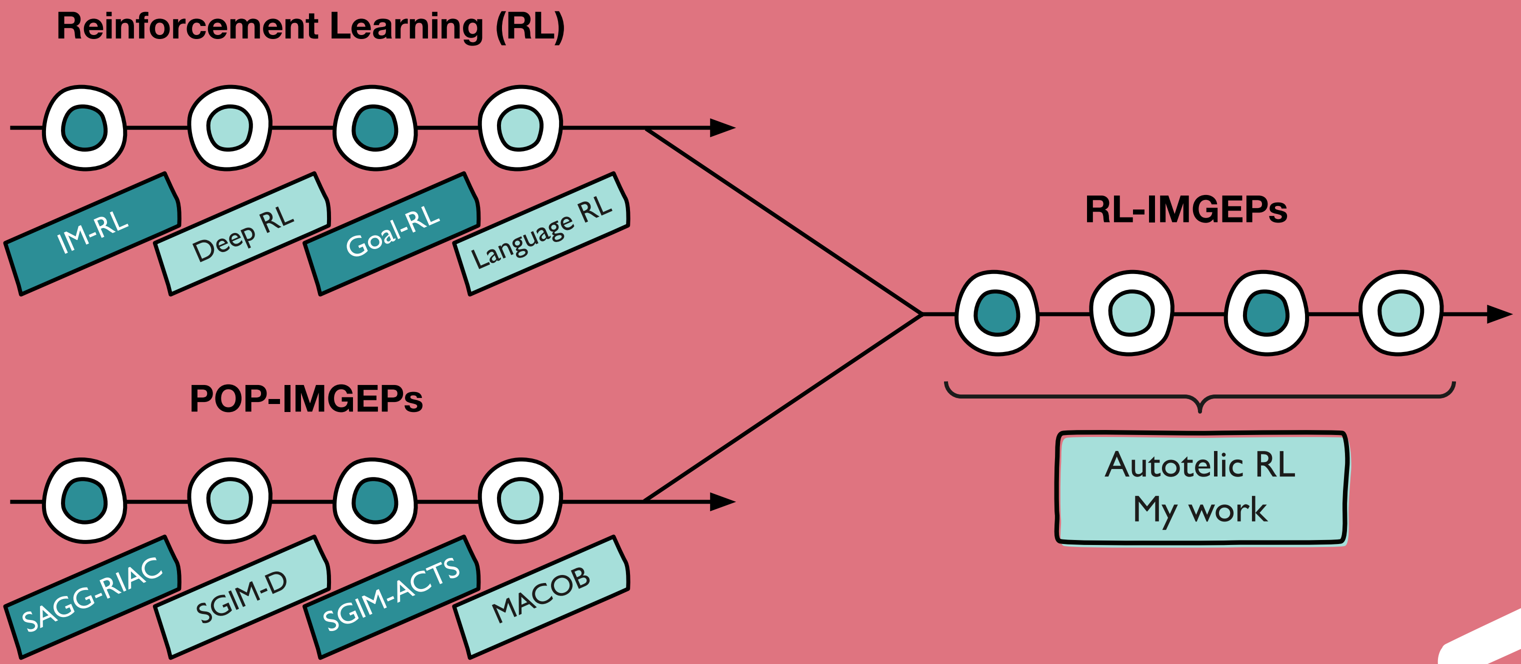
one policy → many goal



+	−
Good finetuning Good generalization Better sample efficiency Low memory (1 policy)	Requires differentiability Low exploration

Towards Autotelic Reinforcement Learning

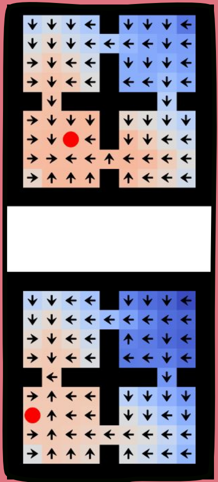
GEP-PG, first combination of POP-IMGEP and RL (DDPG) (Colas et al., 2018)



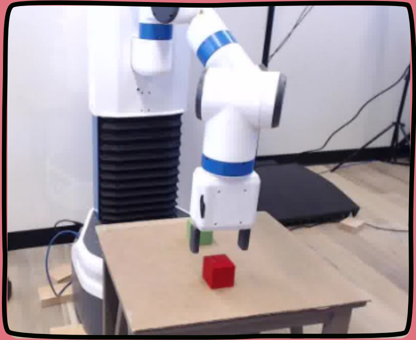
RL (Sutton & Barto, 1998); IM-RL (Schmidhuber, 1991); Goal-RL (Schaul, 2015); Language RL (Herman, 2017)

Goal-Conditioned RL

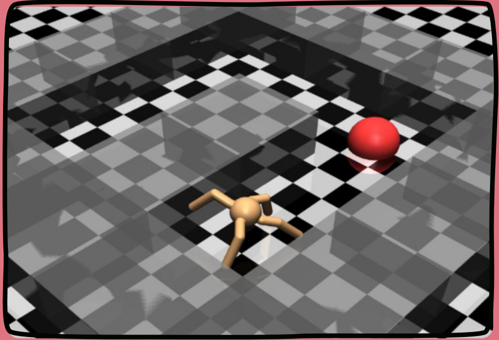
Algorithm families	Learning to represent goals	Learning to sample goals	Learning to evaluate goal-reaching	References
Goal-conditioned RL (GC-RL)	Hand-coded representations ✗	External goals ✗	Hand-coded reward functions ✗	Schaul, 2015; Andrychowicz, 2017
Curriculum GC-RL	Hand-coded representations ✗	Learned sampling	Hand-coded reward functions ✗	Florensa, 2018; Sukhbaatar, 2018; **Colas, 2019**



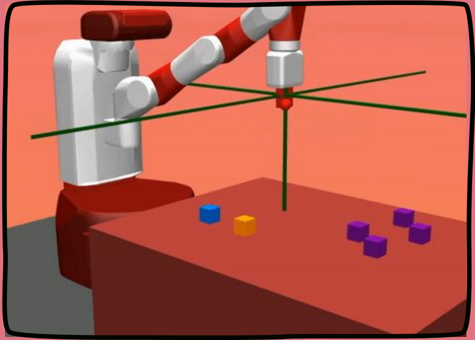
Schaul, 2015



Andrychowicz, 2017



Florensa, 2018

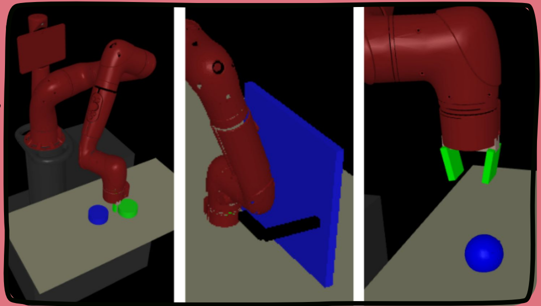


Colas, 2019

Piagetian Autotelic RL

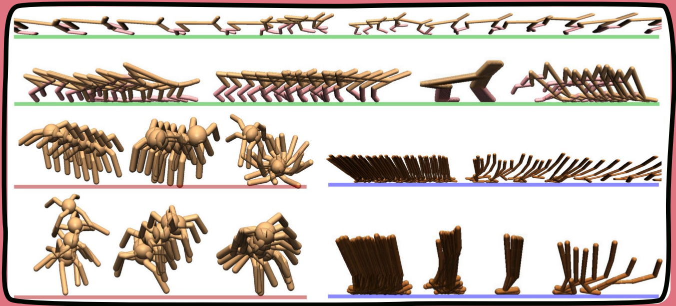
Algorithm families	Learning to represent goals	Learning to sample goals	Learning to evaluate goal-reaching	References
Visual GC-RL	Auto-encoders (visual goals)	Sampling with latent prior	Distance in latent space	Nair, 2018/2020; Pong, 2019; Pitis, 2020
Unsupervised Skill discovery	Categorical + skill discriminability	Sampling with latent prior	Skill discriminability $R_g(s) = \log q_\theta(z_g \mid s)$	Eysenbach, 2018; Sharma, 2020; Campos, 2020

Piagetian Autotelic RL: agents learn goal representations and reward functions on their own through intrinsic motivations, environment interactions and learning.



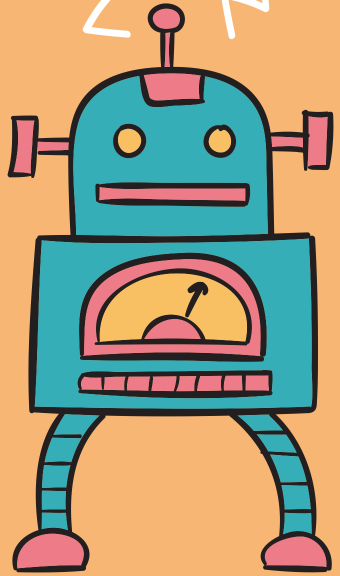
Nair, 2018; Pong, 2019

Eysenbach, 2018



Vygotskian Autotelic Learning

A complementary view on skill learning

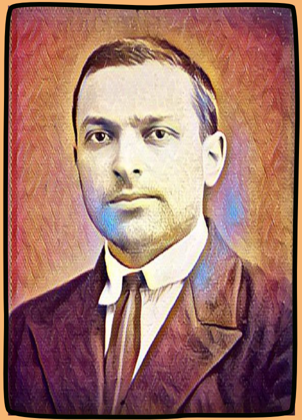


Social Autotelic Learners

Open-ended
repertoire of skills

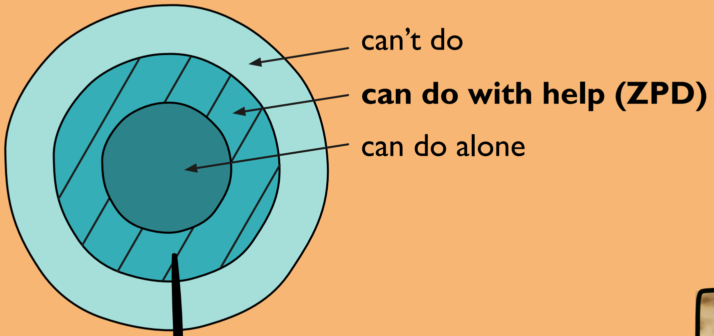
Social Situatedness
Humans learn from others in a rich socio-cultural world.

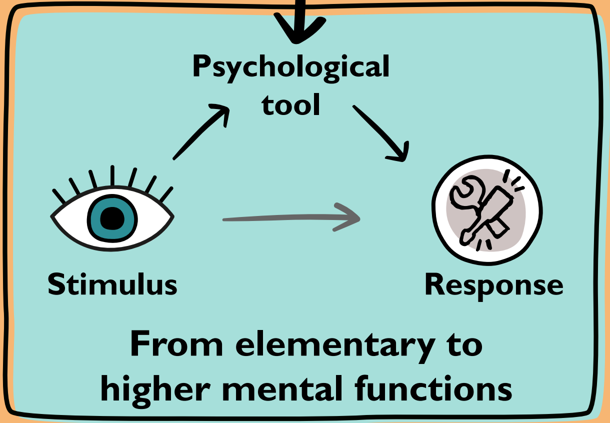
Vygotskian View on Human Development

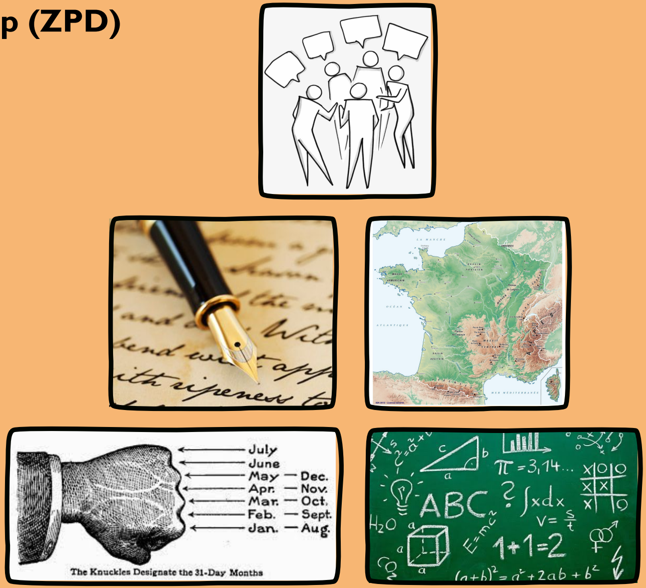


Lev Vygotsky (1896-1934)

Zone of Proximal Development (ZPD)

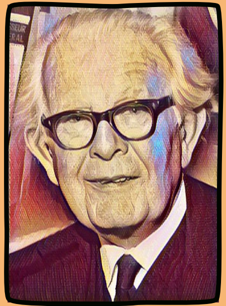






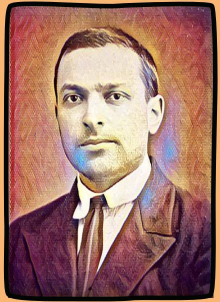
Examples of psychological tool

Language as a Cognitive Tool



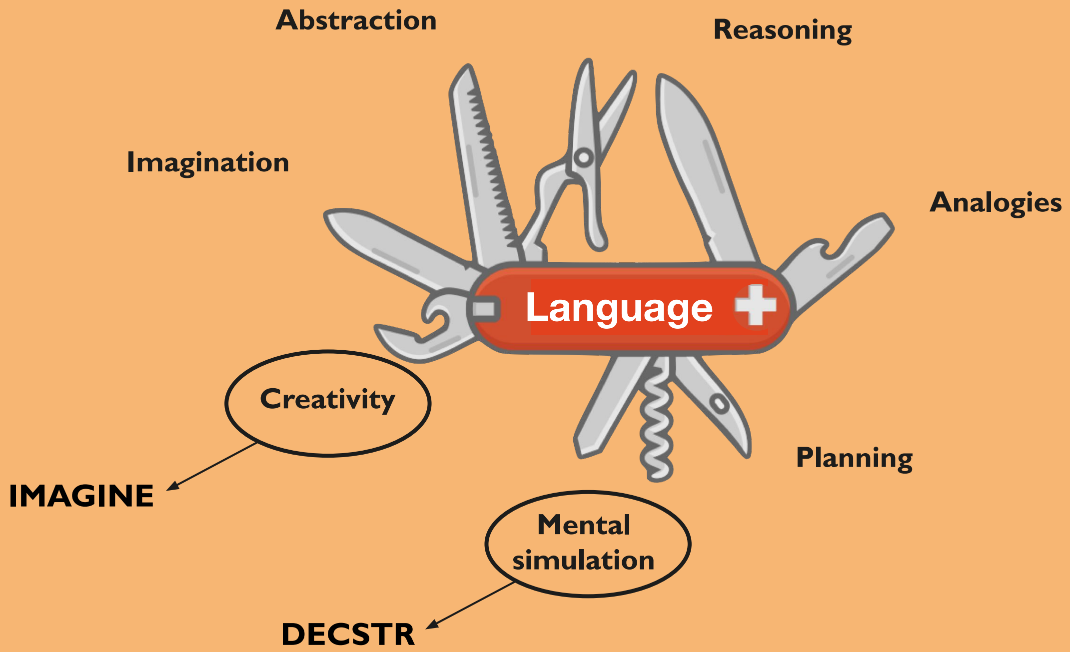
Jean Piaget

Egocentric speech is a sign of cognitive immaturity.
inside → outside



Lev Vygotsky

Egocentric speech is the internalization of social speech.
outside → inside



Vygotsky, 1934; Berk, 1994; Gentner, 1983/2017; Clark 1998; Hermer-Vazquez, 2001; Carruthers, 2002; Lupyan, 2012; Bergen, 2012.

Vygotskian Autotelic Agents

Algorithm families	Learning to represent goals	Learning to sample goals	Learning to evaluate goal-reaching	References
Linguistic GC-RL	Learned linguistic representations	External goals ✗	External rewards / Hand-coded reward functions ✗	Hermann, 2017; Chan, 2019; Jiang, 2019; Chevalier-Boisvert, 2019; Côté, 2019; Hill, 2020.

Vygotskian Autotelic Agent

Internalize cultural goal representations, goal selection and biases.

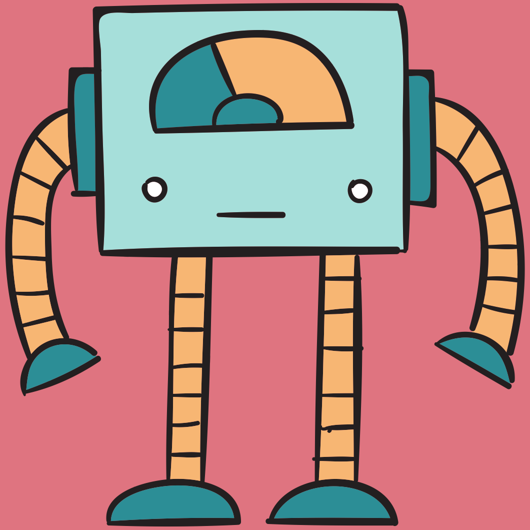
1. to learn autonomously
2. to perform structured exploration
3. to use language as a cognitive tool

Evaluation of Autotelic Agents

- **Measure exploration**
  (coverage, interesting interactions)
- **Measure generalization**
  (performance on held-out set of goals)
- **Measure transfer learning**
  (downstream tasks: fine-tuning, hierarchical setting)
- **Measure robustness**
  (non-controllable objects, nonstationarities)
- **Open the black box**
  (learned representations, learned goal sampler, developmental trajectories)

IMAGINE

Linguistic creativity for exploration and generalization.



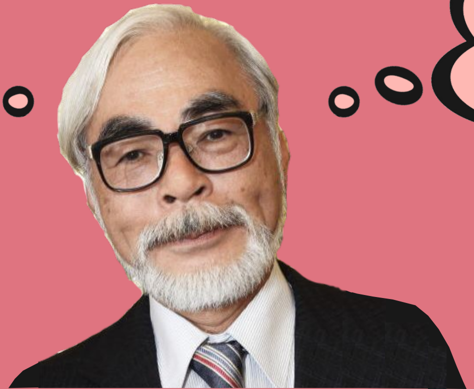
Towards Out-of-Distribution Goal Generation

Linguistic Creativity

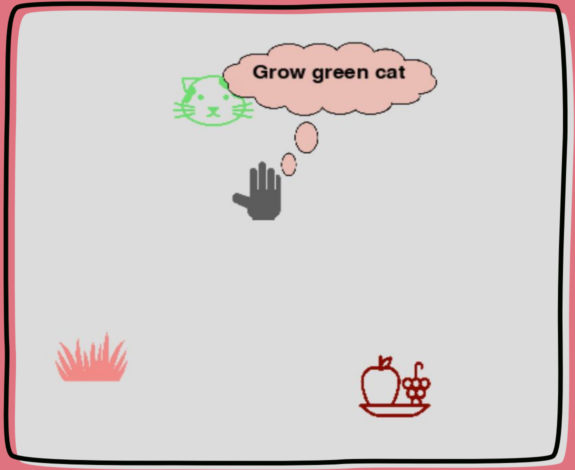
Creativity = novelty x appropriateness.
(Simonton, 2012)

Linguistic creativity: generate new utterances (novelty) from a known grammar/known constructions (appropriateness).
(Chomsky, 1957; Hoffmann, 2020)

cat + bus
=
cat-bus!



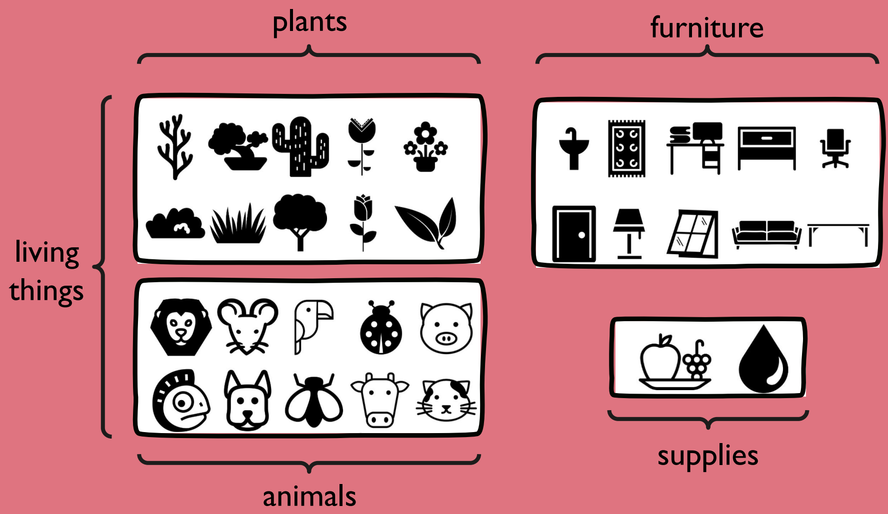
Playground Environment



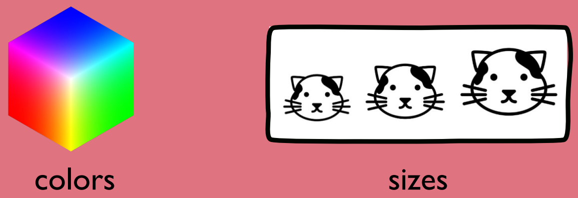
Social descriptions

- Go (e.g. "go top left")
- Grasp (e.g. "grasp red plant"),
- Grow (e.g. "grow any lion").

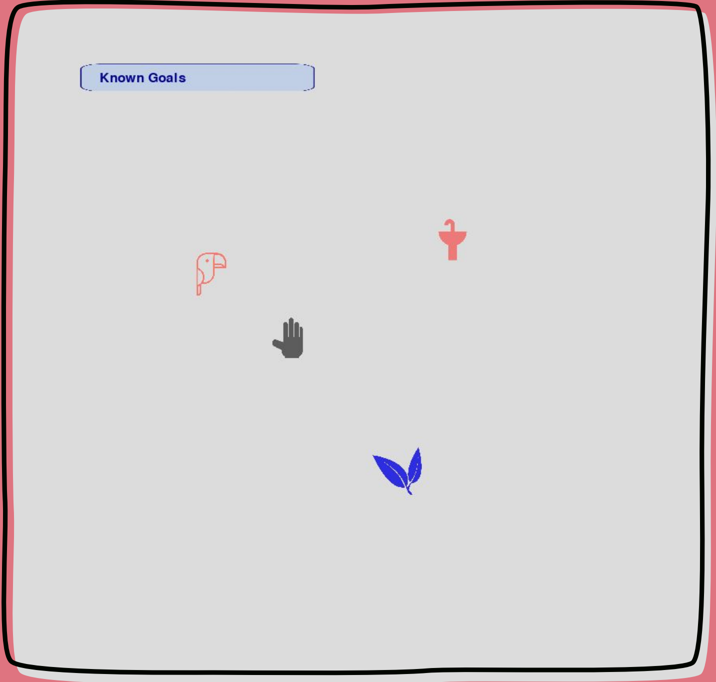
Objects:



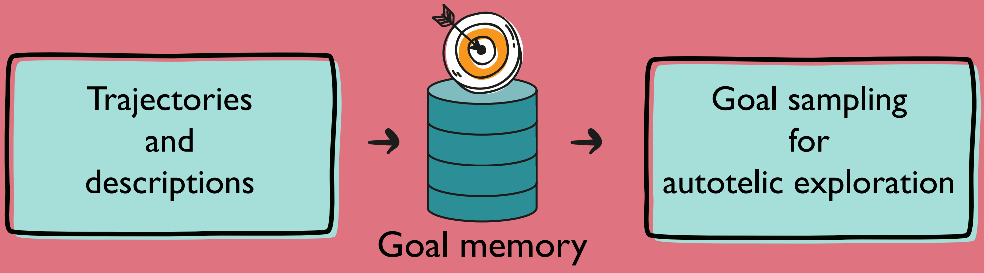
Procedural generation:

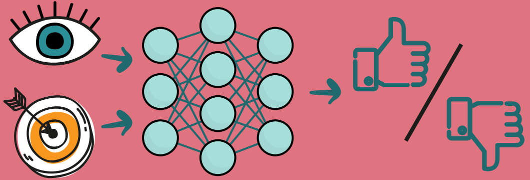


Internalization of Linguistic Goals



Autotelic exploration with social descriptions



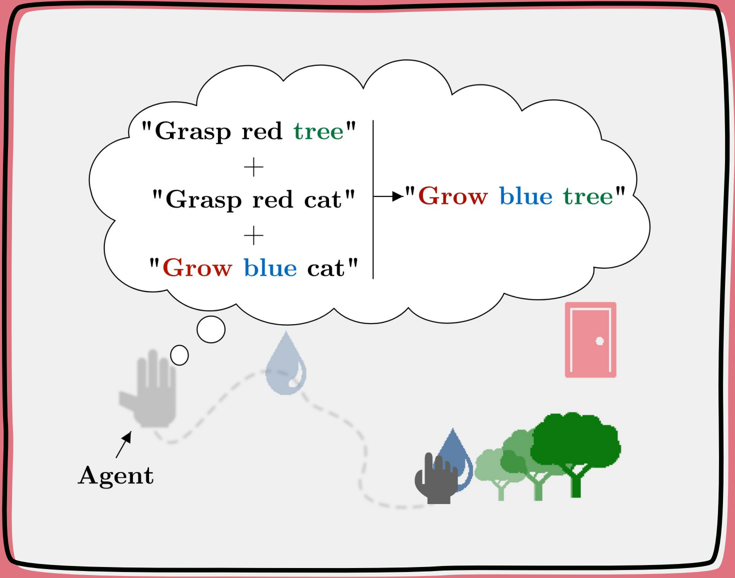


$$R_g(s,\, a) = r$$

Learn a goal-conditioned reward function

(Bahdanau, 2019)

Language as a Cognitive Tool to Imagine Goals



Creative autotelic exploration

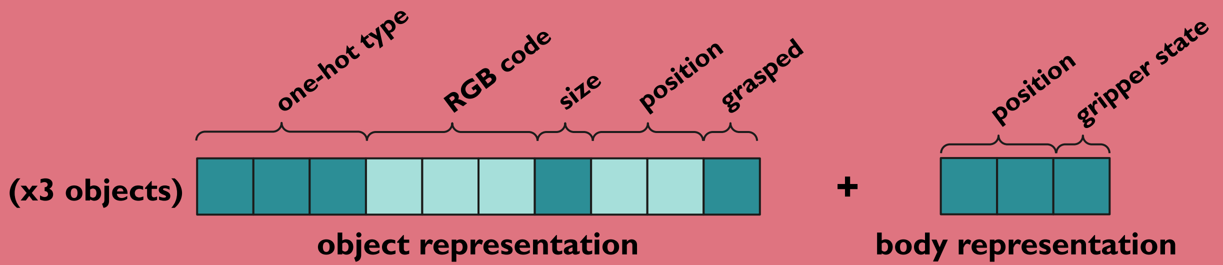
Idea

Use language compositionality to systematically compose novel, out-of-distribution goals.

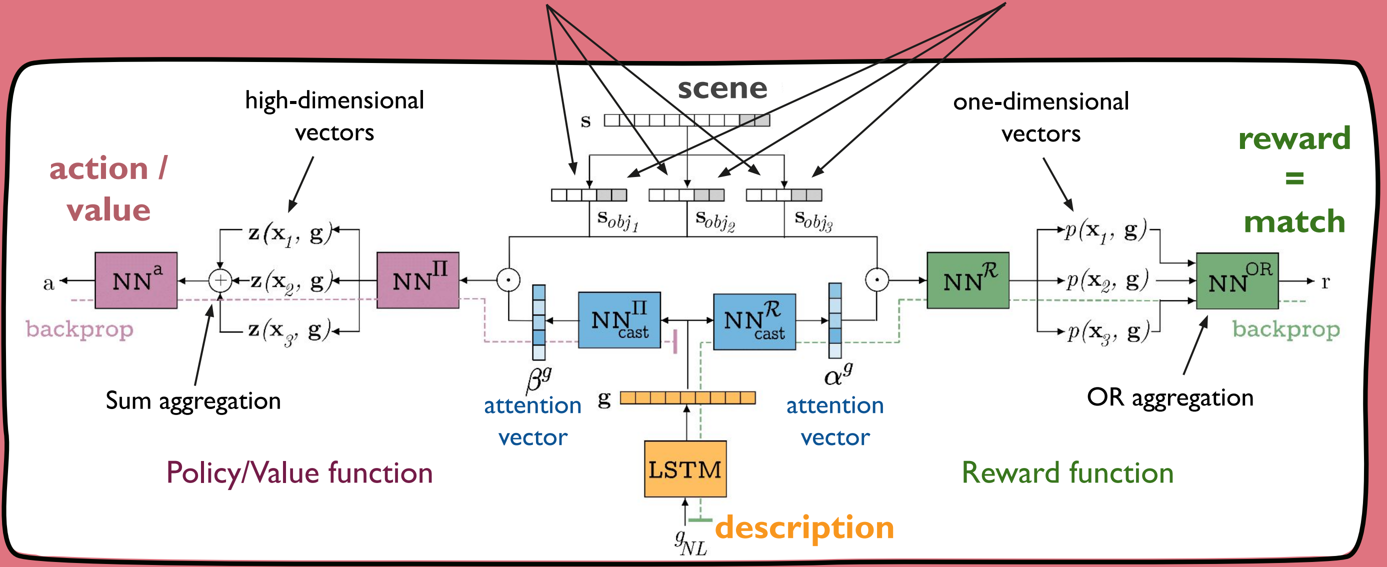
Internalized goal generation and reward functions let the agent train autonomously.

Object-Centered Inductive Biases

Object-Based Representations
(Spelke, 2003; Mandler, 1999)



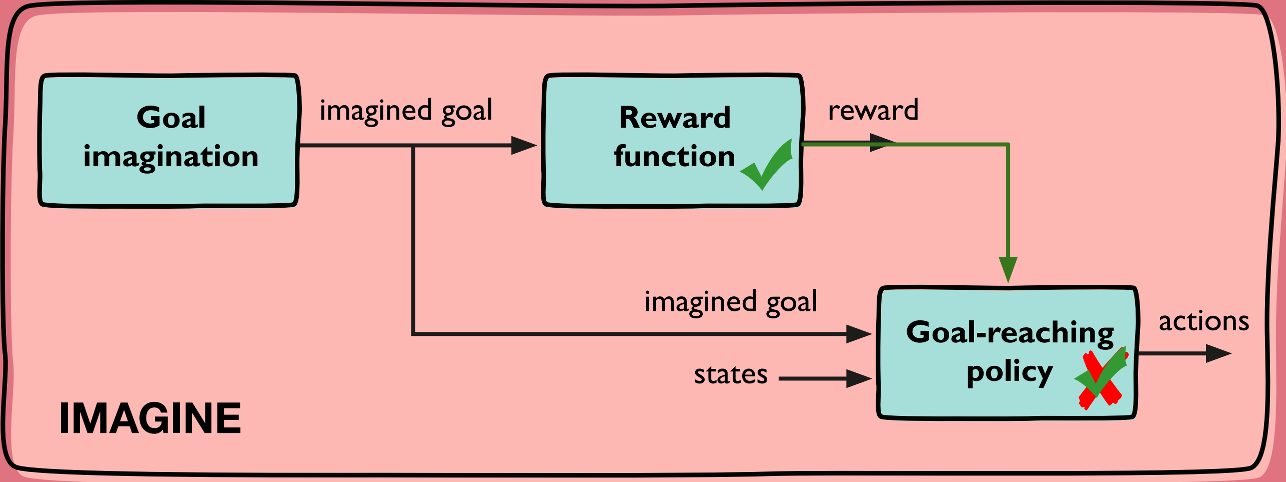
Deep-Sets Architectures
(Zaheer, 2017)



Testing Systematic Generalization

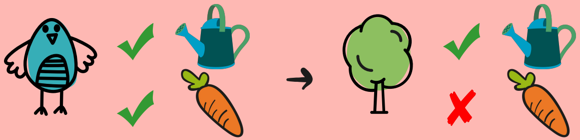
Several types of generalization

- **Zero-shot policy**: the agent can reach imagined goals.
- **Zero-shot reward function**: the agent can recognize matching scenes for new goals.



Testing Systematic Generalization

Enhancing generalization with goal imagination

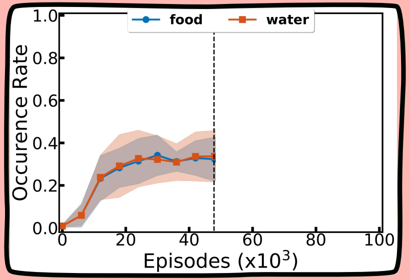


Train set

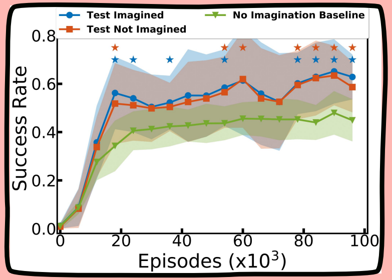
Test set

Agents correct for overgeneralizations of their policy thanks to their reward function.

This effects transfers to similar, non-imagined goals.



Behavioral adaptation



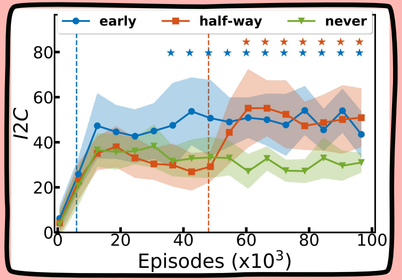
Beyond imagination generalization

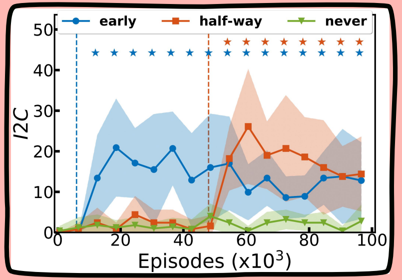
Effects of a Creative Autotelic Exploration

Effect on exploration

Social goals are biased towards objects and interactions.

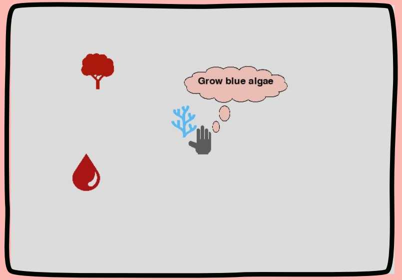
Imagined goals are similarly biased and creative: they drive agents to explore their world.

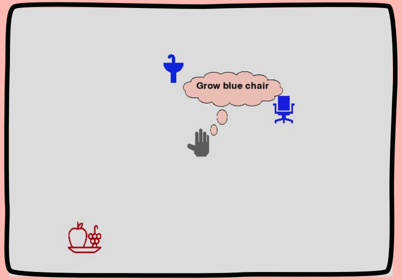




Feeding plants

Feeding furniture





Discussion

Simple mechanism for enhanced generalization

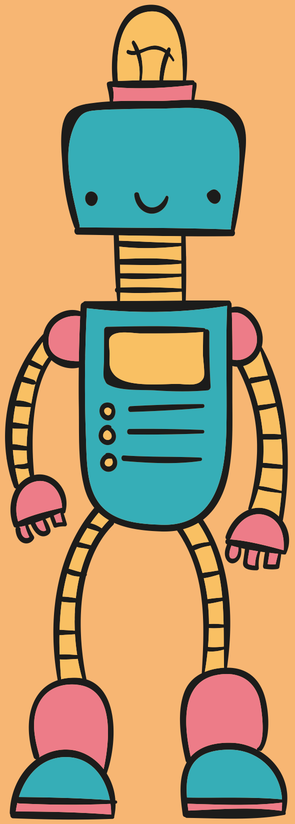
Internalization of reward function + systematic goal imagination let agents fine-tune their policy and generalize better. See also *Good-Enough Compositional Data Augmentation (Andreas, 2019).*

Distinctively human play

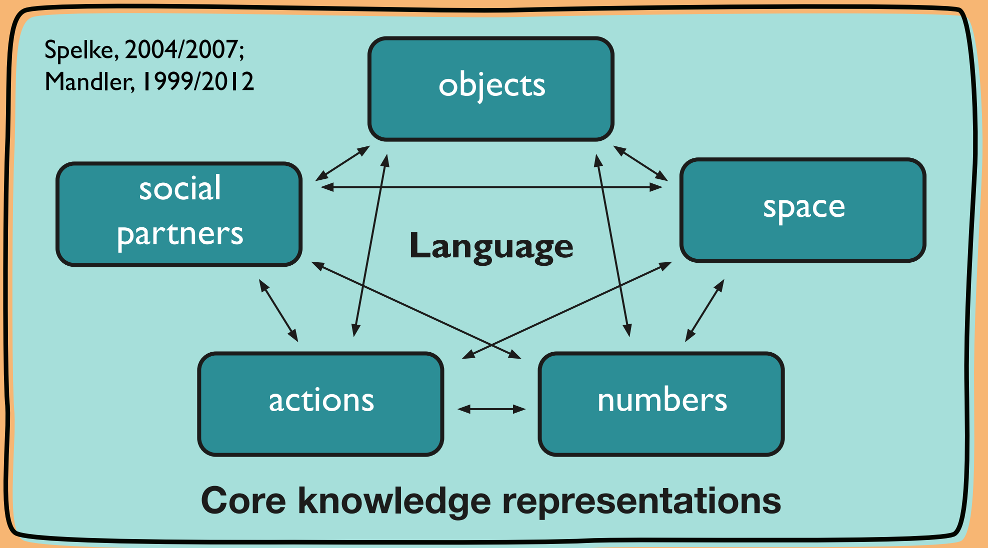
Children set arbitrary goals to themselves during pretend play. Problems constrain the search for hypotheses and plans such that children learn to efficiently generate new hypotheses by training on arbitrary problems (Chu & Schulz, 2020a,b).

DECSTR

Mental simulation of possible futures,
a new form of language grounding.



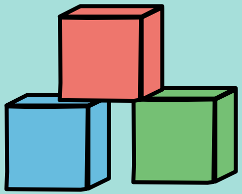
Motivations



How to combine concrete goals of preverbal infants with abstract linguistic goals?

Core Knowledge of Objects

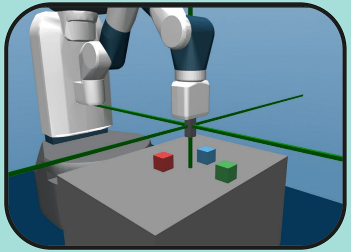
Assume semantic spatial predicates



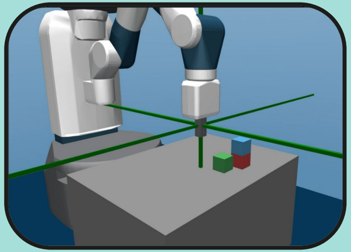
Semantic configuration

1	1	1	1	0	1	0	0	0

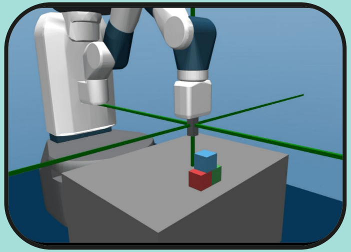
close predicates above predicates



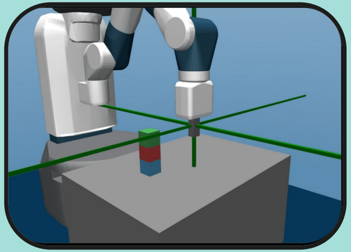
000000000



111000100

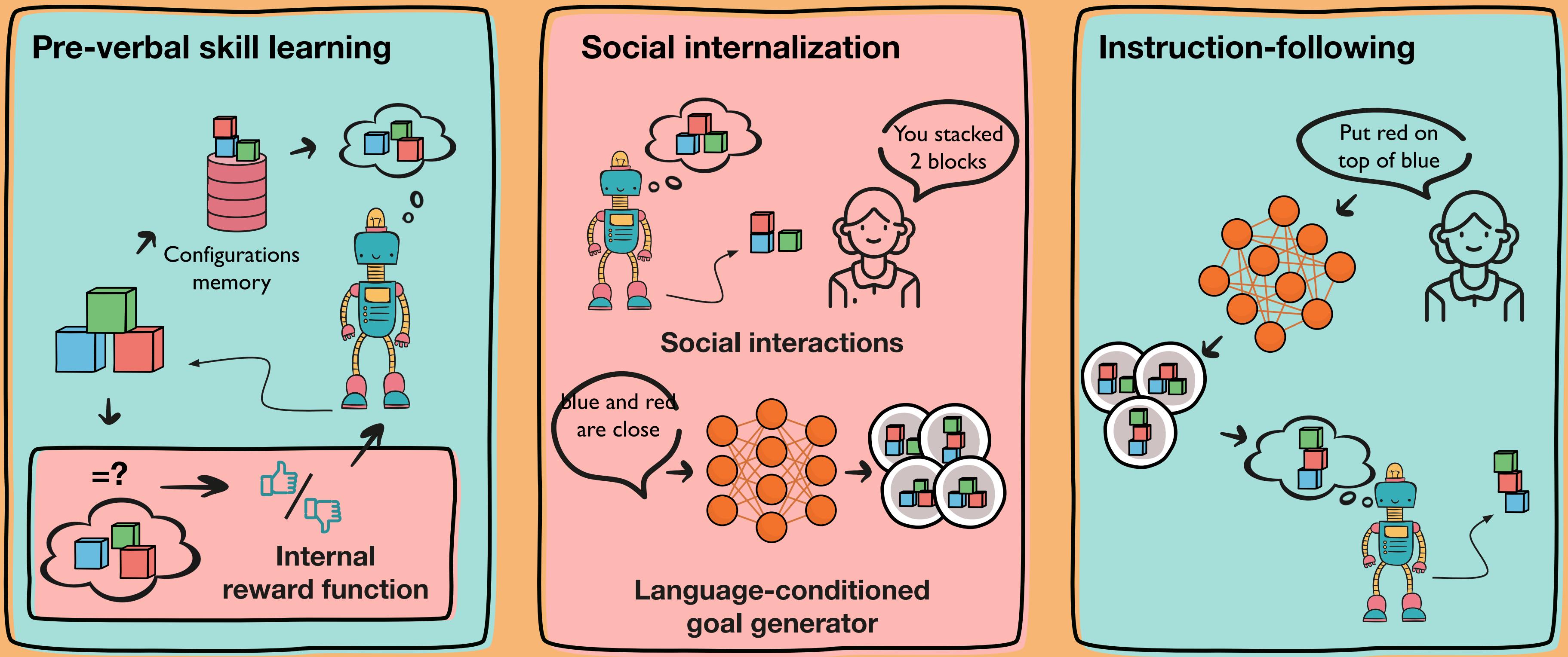


111000101

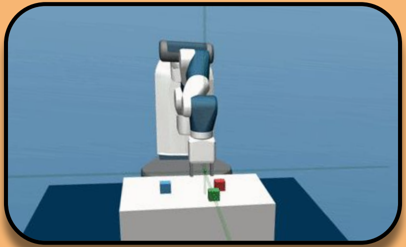


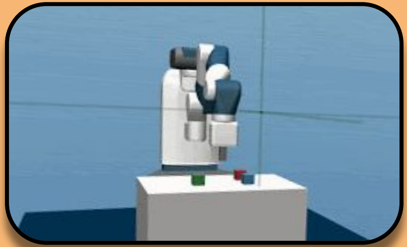
111010100

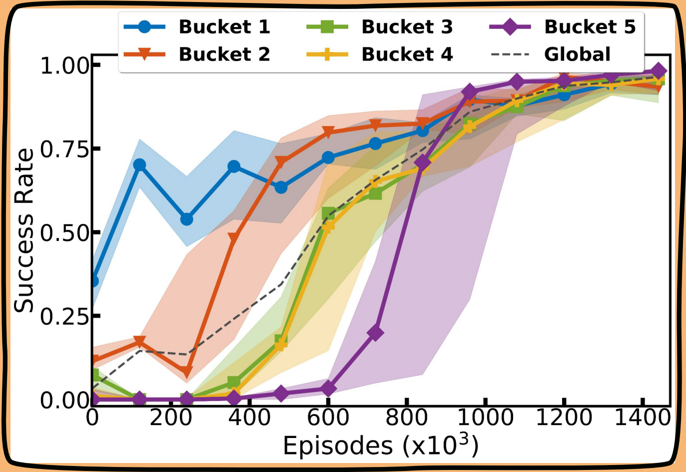
Grounding Language in Goals

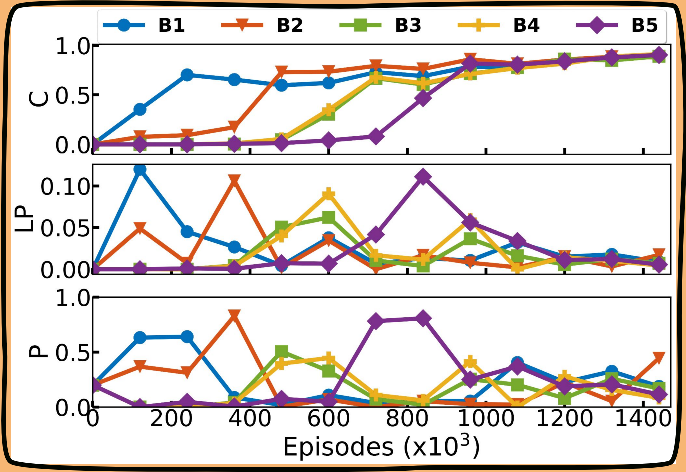


Preverbal Skill Learning (Phase 1)

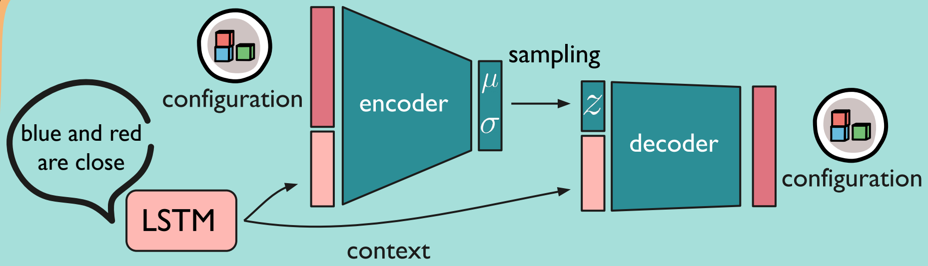








Social Internalization (Phase 2)



$loss = \mathcal{D}_{KL}\left[q(z \mid x, c) \mid\mid \mathcal{N}(0,\, 1)\right] + \text{cross-entropy}(\ ,\ )$

Conditional Variational Auto-Encoder

Examples of descriptions:
"you put red below green"
"you got blue and green close"

and more abstract:
"you built a pyramid"
"you made a construction"
"you got green on top"

We tested:

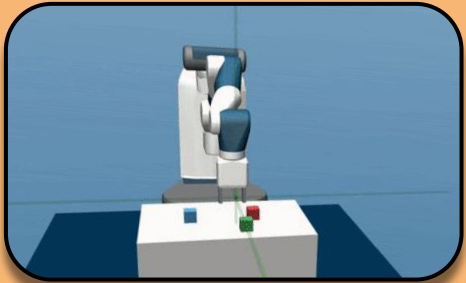
- precision (valid configurations)
- recall (diversity).

Results:

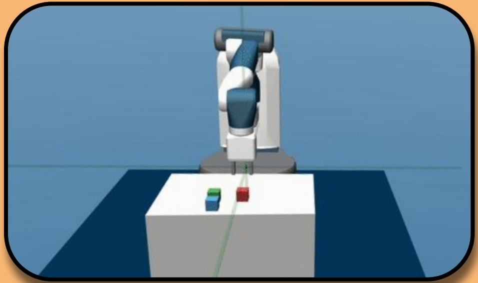
- Near-perfect scores on the train set,
- Generalizes systematically to new sentences,
- Did not work on continuous state generation.

(C-VAE: Sohn, 2015)

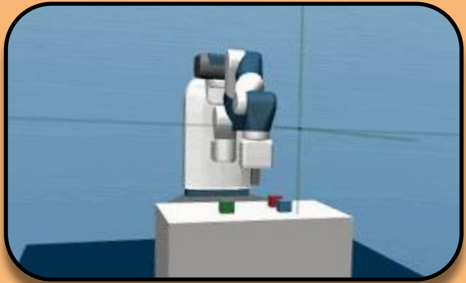
Instruction-Following (Phase 3)



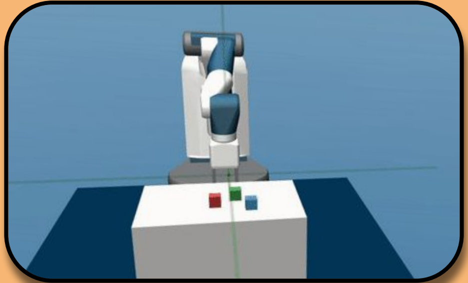
"Get green close to blue"



"Make a pyramid"

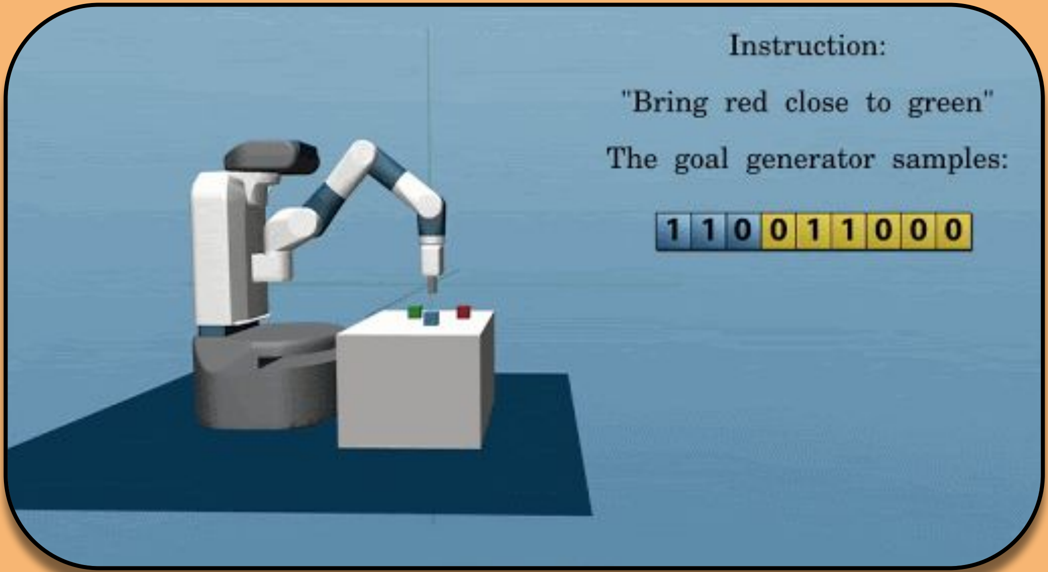


"Put green below red"

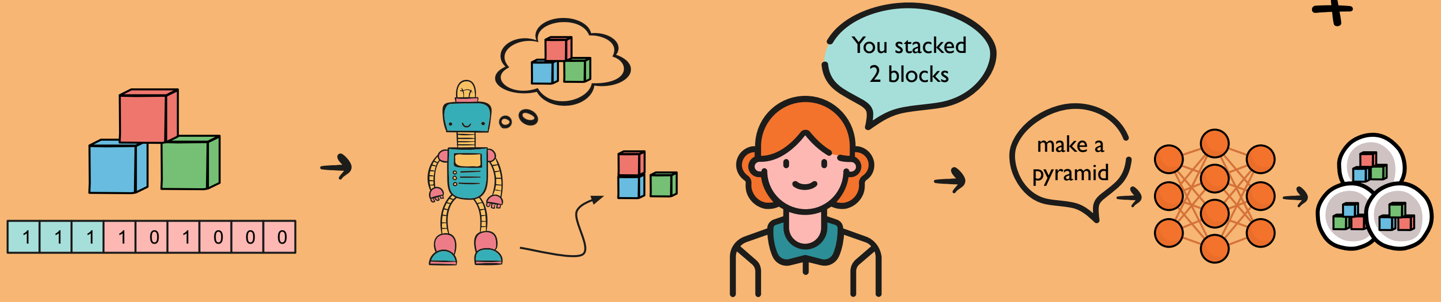


"Build a stack"

Strategy-Switching

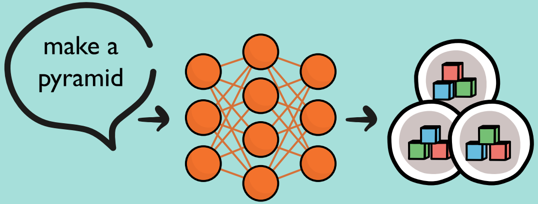


Discussion



Abstraction

Grounding language in core knowledge offers abstraction, categorization by examples.



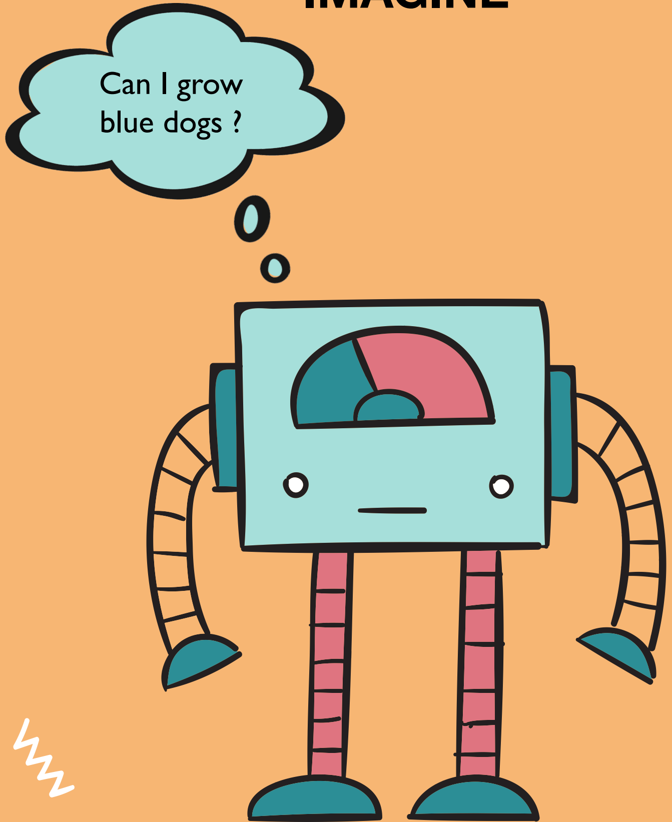
Abstract semantic predicates

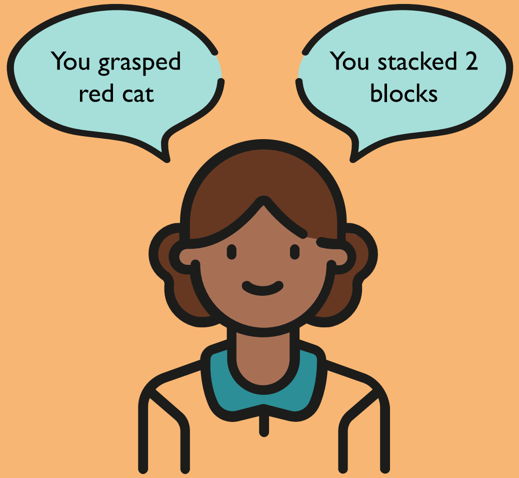
Linguistic statements can be new predicates.

e.g.
“door is open”
“object is red”

Vygotskian Autotelic Agents

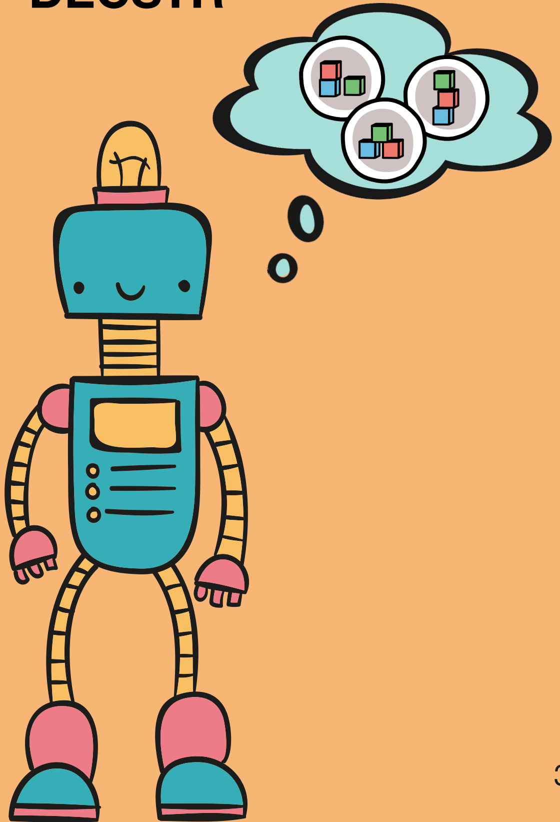
IMAGINE





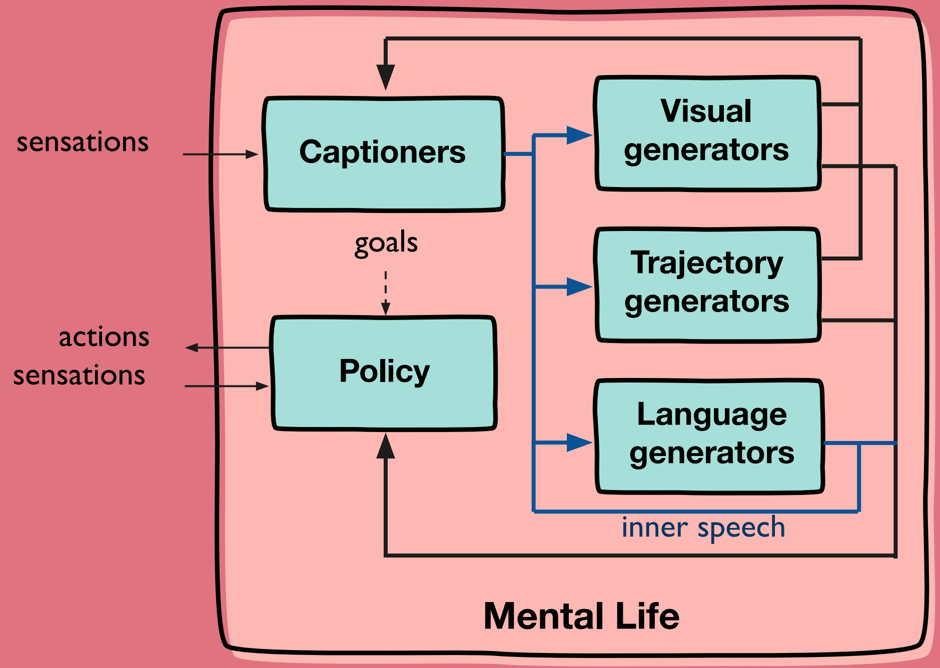
Social Partner

DECSTR



Perspectives

Artificial Mental Life



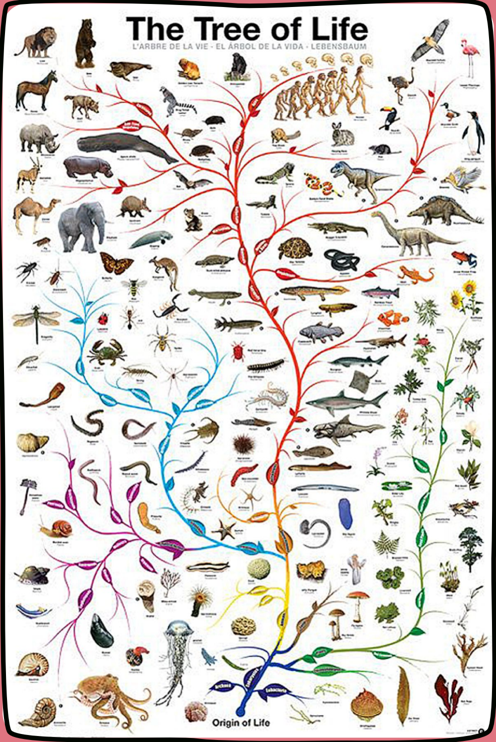
Open-Endedness

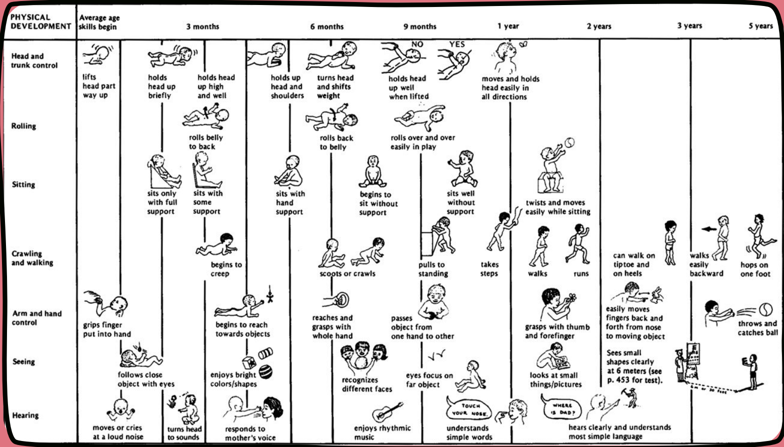
Open-ended processes

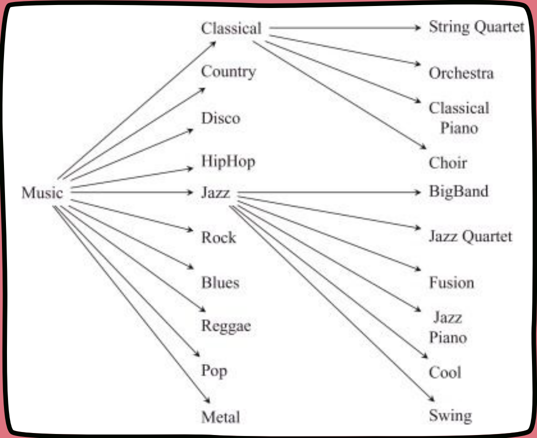
- Natural evolution
- Music
- Science
- Human skill learning
- …

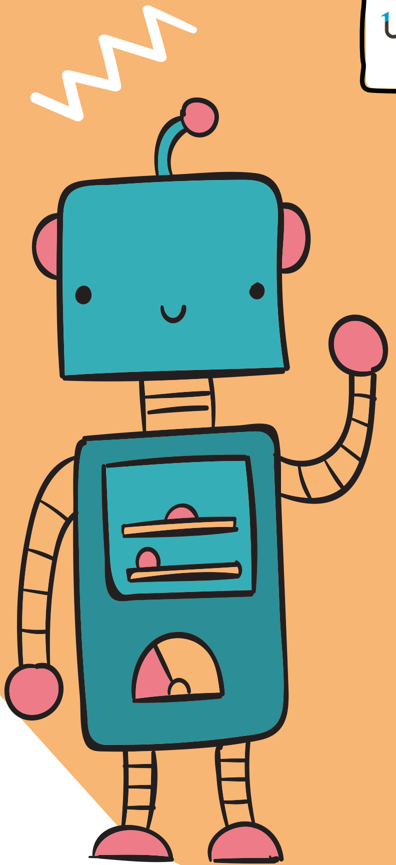
How to evaluate open-ended remains an open question (Hintze, 2019; Stanley, 2019).

Why not integrate agents into *our* open-ended cultural world? It seems to work for children.









Thanks!

Supervisors:

Pierre-Yves Oudeyer
(INRIA)

Olivier Sigaud
(Sorbonne Univ)

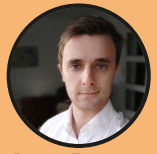
Collaborators:

Ahmed Akazia
(Sorbonne Univ)

Mohamed Chetouani
(Sorbonne Univ)

Jeff Clune
(OpenAI)

Peter Ford Dominey
(INSERM)



Pierre Fournier
(ex Sorbonne)

Boris Hejblum
(INRIA)

Katja Hofmann
(MSR)

Joost Huizinga
(OpenAI)

Tristan Karch
(INRIA)

Nicolas Lair
(INSERM)

Vashisht Madhavan
(ex UberAI)

Clément
Moulin-Frier (INRIA)

Rémy Portelas
(INRIA)

Mélanie Prague
(INRIA)

Sébastien Rouillon
(Bordeaux Univ.)

Rodolphe
Thiebaut (INRIA)

Lilian Weng
(OpenAI)

Jury:

Marc Bellemare
(McGill, Google Brain)

Felix Hill
(DeepMind)

Mehdi Khamassi
(CNRS)

Sebastian Risi
(IT Copenhagen)

Laura Schulz
(MIT)

Harm van Seijen
(Microsoft Research)

Contact:
cedric.colas@inria.fr
https://ccolas.github.io/
https://github.com/flowersteam

References

Andreas, Jacob. "Good-Enough Compositional Data Augmentation." *2020.*

Andrychowicz, Marcin, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. "Hindsight Experience Replay." 2017.

Baranes, Adrien, and Pierre-Yves Oudeyer. "Active Learning of Inverse Models with Intrinsically Motivated Goal Exploration in Robots." *2013.*

———. "Intrinsically Motivated Goal Exploration for Active Motor Learning in Robots: A Case Study." 2010.

Bellemare, Marc, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Remi Munos. "Unifying Count-Based Exploration and Intrinsic Motivation." 2016.

Bergen, Benjamin K. *Louder Than Words: The New Science of How the Mind Makes Meaning*. 2012.

Berk, Laura E. "Why Children Talk to Themselves." *1994.*

Berlyne, Daniel E. "Novelty and Curiosity as Determinants of Exploratory Behaviour." *1950.*

Burda, Yuri, Harrison Edwards, Amos Storkey, and Oleg Klimov. "Exploration by Random Network Distillation." *2018.*

Campos, Víctor, Alexander Trott, Caiming Xiong, Richard Socher, Xavier Giro-i-Nieto, and Jordi Torres. "Explore, Discover and Learn: Unsupervised Discovery of State-Covering Skills." *2020.*

Carruthers, Peter. "Modularity, Language, and the Flexibility of Thought." *2002.*

Chan, Harris, Yuhuai Wu, Jamie Kiros, Sanja Fidler, and Jimmy Ba. "ACTRCE: Augmenting Experience via Teacher's Advice For Multi-Goal Reinforcement Learning." *2019.*

Chevalier-Boisvert, Maxime, Dzmitry Bahdanau, Salem Lahlou, Lucas Willems, Chitwan Saharia, Thien Huu Nguyen, and Yoshua Bengio. "Baby-Ai: First Steps Towards Grounded Language Learning with a Human in the Loop." 2019.

Chomsky, Noam. *Syntactic Structures*. 1957.

Chu, Junyi, and Laura Schulz. "Exploratory Play, Rational Action, and Efficient Search." 2020.

Chu, Junyi, and Laura E. Schulz. "Play, Curiosity, and Cognition." *2020.*

Clark, Andy. "Magic Words: How Language Augments Human Computation." 1998.

Côté, Marc-Alexandre, Akos Kadar, Xingdi Yuan, Ben Kybartas, Tavian Barnes, Emery Fine, James Moore, et al. "TextWorld: A Learning Environment for Text-Based Games." 2018.

Czikszentmihalyi, Mihaly. *Flow: The Psychology of Optimal Experience*. 1990.

Eysenbach, Benjamin, Abhishek Gupta, Julian Ibarz, and Sergey Levine. "Diversity Is All You Need: Learning Skills Without a Reward Function." 2018.

Florensa, Carlos, David Held, Xinyang Geng, and Pieter Abbeel. "Automatic Goal Generation for Reinforcement Learning Agents." 2018.

Forestier, Sébastien, and Pierre-Yves Oudeyer. "Modular Active Curiosity-Driven Discovery of Tool Use." 2016.

Gentner, Dedre. "Structure-Mapping: A Theoretical Framework for Analogy." *1983.*

Gentner, Dedre, and Christian Hoyos. "Analogy and Abstraction." *2017.*

Hermann, Karl Moritz, Felix Hill, Simon Green, Fumin Wang, Ryan Faulkner, Hubert Soyer, David Szepesvari, et al. "Grounded Language Learning in a Simulated 3D World." *2017.*

Hermer-Vazquez, L. "Language, Space, and the Development of Cognitive Flexibility in Humans: The Case of Two Spatial Memory Tasks." *2001.*

Hill, Felix, Andrew Lampinen, Rosalia Schneider, Stephen Clark, Matthew Botvinick, James L. McClelland, and Adam Santoro. "Environmental Drivers of Systematicity and Generalization in a Situated Agent." *2020.*

Hoffmann, Thomas. "Construction Grammar and Creativity: Evolution, Psychology, and Cognitive Science." *2020.*

References

Kaplan, Frédéric, and Pierre-Yves Oudeyer. “Maximizing Learning Progress: An Internal Reward System for Development.” 2004.

Laversanne-Finot, Adrien, Alexandre Péré, and Pierre-Yves Oudeyer. “Curiosity Driven Exploration of Learned Disentangled Goal Spaces.” 2018.

Lupyan, Gary. “What Do Words Do? Toward a Theory of Language-Augmented Thought.” 2012.

Mandler, Jean M. “On the Spatial Foundations of the Conceptual System and Its Enrichment.” *2012.*

Mandler, Jean M. “Preverbal Representation and Language.” *1999.*

———. “Thought Before Language.” 2004.

Moulin-Frier, Clément, Sao Mai Nguyen, and Pierre-Yves Oudeyer. “Self-Organization of Early Vocal Development in Infants and Machines: The Role of Intrinsic Motivation.” *2014.*

Nair, Ashvin, Shikhar Bahl, Alexander Khazatsky, Vitchyr Pong, Glen Berseth, and Sergey Levine. “Contextual Imagined Goals for Self-Supervised Robotic Learning.” 2020.

Nair, Ashvin V, Vitchyr Pong, Murtaza Dalal, Shikhar Bahl, Steven Lin, and Sergey Levine. “Visual Reinforcement Learning with Imagined Goals.” 2018.

Nguyen, Sao Mai, Adrien Baranes, and Pierre-Yves Oudeyer. “Bootstrapping Intrinsically Motivated Learning with Human Demonstrations.” *2011.*

Nguyen, Sao Mai, and Pierre-Yves Oudeyer. “Active Choice of Teachers, Learning Strategies and Goals for a Socially Guided Intrinsic Motivation Learner.” *2012.*

Oudeyer, Pierre-Yves, and Frederic Kaplan. “What Is Intrinsic Motivation? A Typology of Computational Approaches.” *2007*

Pathak, Deepak, Pulkit Agrawal, Alexei A Efros, and Trevor Darrell. “Curiosity-Driven Exploration by Self-Supervised Prediction.” 2017.

Pere, Alexandre, Sebastien Forestier, Olivier Sigaud, and Pierre-Yves Oudeyer. “Unsupervised Learning of Goal Spaces for Intrinsically Motivated Goal Exploration.” 2018.

Pitis, Silviu, Harris Chan, Stephen Zhao, Bradly Stadie, and Jimmy Ba. “Maximum Entropy Gain Exploration for Long Horizon Multi-Goal Reinforcement Learning.” 2020.

Pong, Vitchyr H, Murtaza Dalal, Steven Lin, Ashvin Nair, Shikhar Bahl, and Sergey Levine. “Skew-Fit: State-Covering Self-Supervised Reinforcement Learning.” 2019.

Schaul, Tom, Daniel Horgan, Karol Gregor, and David Silver. “Universal Value Function Approximators.” 2015.

Schmidhuber, Jürgen. “A Possibility for Implementing Curiosity and Boredom in Model-Building Neural Controllers.” 1991

Sharma, Archit, Shixiang Gu, Sergey Levine, Vikash Kumar, and Karol Hausman. “Dynamics-Aware Unsupervised Discovery of Skills.” *2020.*

Simonton, Dean Keith. “Creative Productivity and Aging.” 2012.

Sohn, Kihyuk, Honglak Lee, and Xinchen Yan. “Learning Structured Output Representation Using Deep Conditional Generative Models.” 2015.

Spelke, Elizabeth S, ed. *What Makes Us Smart*. 2003.

Spelke, Elizabeth S., and Katherine D. Kinzler. “Core Knowledge.” *2007.*

Steels, Luc. “The Autotelic Principle.” 2004.

Sukhbaatar, Sainbayar, Zeming Lin, Ilya Kostrikov, Gabriel Synnaeve, Arthur Szlam, and Rob Fergus. “Intrinsic Motivation and Automatic Curricula Via Asymmetric Self-Play.” 2017.

Sutton, Richard S, and Andrew G Barto. *Introduction to Reinforcement Learning*. 1998.

Vygotsky, Lev S. *Thought and Language*. 1934.

Zaheer, Manzil, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Russ R Salakhutdinov, and Alexander J Smola. “Deep Sets.” 2017.

Thanks!

Does anyone have any questions?
addyouremail@freepik.com
620 421 838

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, infographics & images by **Freepik**

Please keep this slide for attribution