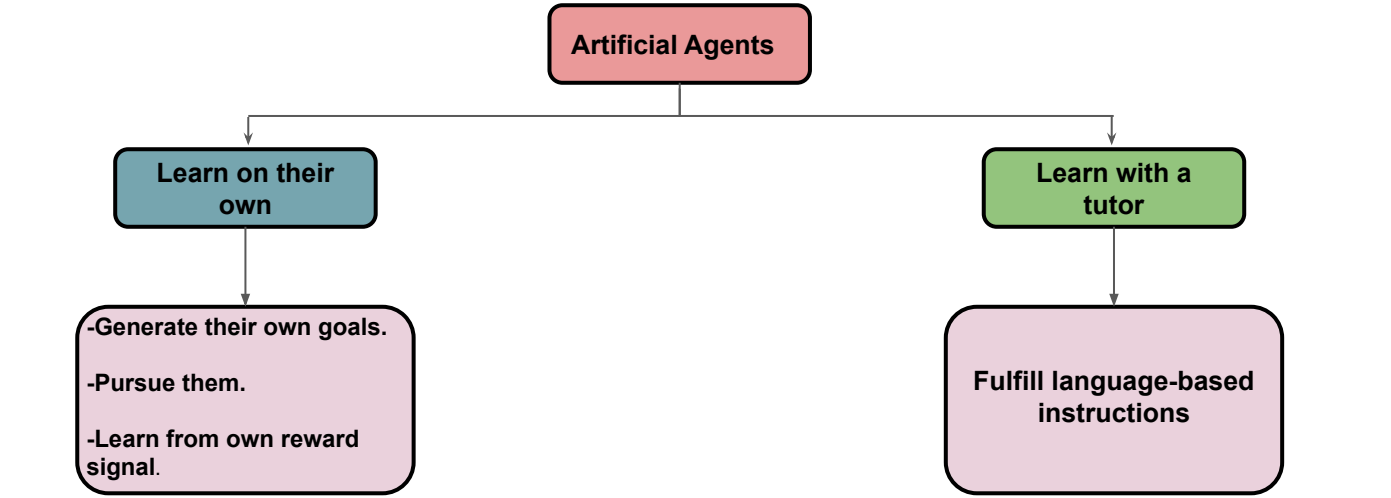
Grounding Language to Autonomously Acquired Skills via Goal Generation
ICLR 2021

Ahmed Akakzia*, **Cédric Colas***, Pierre-Yves Oudeyer, Mohamed Chetouani, Olivier Sigaud

Paris Sorbonne University - INRIA Bordeaux

Contact: ahmed.akakzia@isir.upmc.fr

Motivation



IMGEP[1] family (Forestier et al., 2017)

Instruction-following (Luketina et al. 2010)

It's good to have

Behavioral diversity
Learn different ways to reach a goal

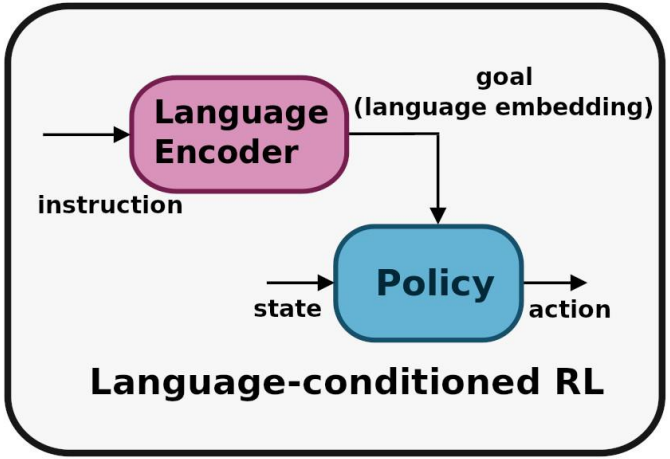
Language-based goals
Respect set of constraints on the state space

[1]Intrinsically Motivated Goal Exploration Processes

Motivation

A viable way to train instruction-following agents:

Language-Conditioned RL (LC RL)



However

- LC RL agents require **language inputs** to behave and learn.

preverbal infants don't !
(Mandler et al., 1978)

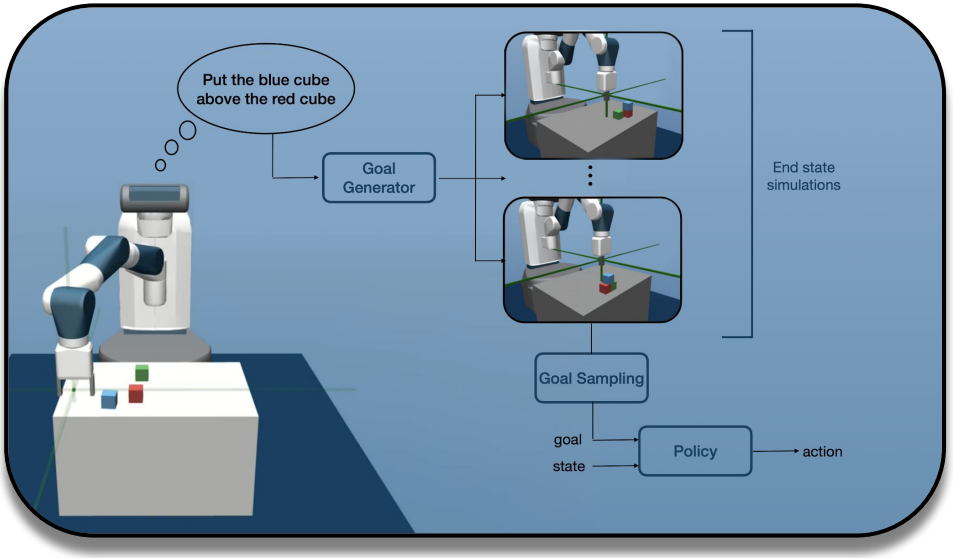
- LC RL limits **behavioral diversity**: there is one way to achieve one instruction.

A Novel architecture

- The **Language-Goal-Behavior** architecture (LGB) to decouple skill learning from language grounding via an **intermediate semantic goal representation**.

- DECSTR[2], an implementation of LGB which operates through 3 phases:

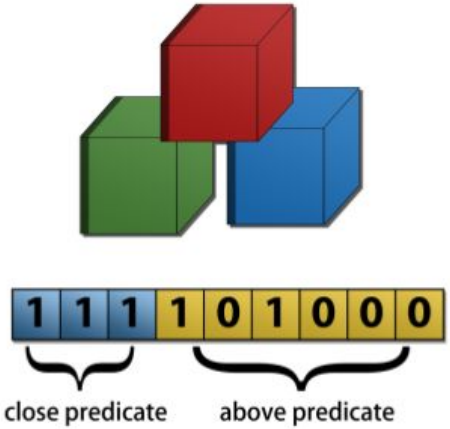
  - Skill learning phase **(G→B)**.
  - Language grounding phase **(L→G)**.
  - Instruction following phase **(L→G→B)**.

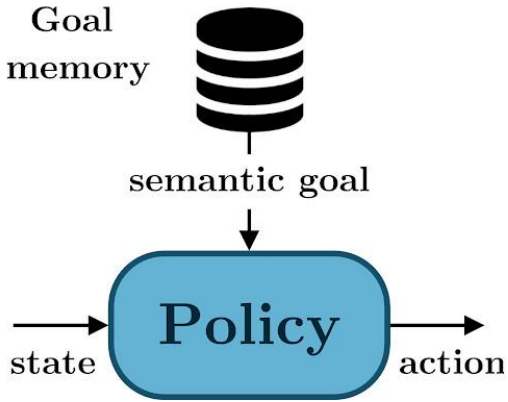


DECSTR shows good behavioral diversity !

[2]**DE**ep sets and **C**urriculum with **S**eman**T**ic goal **R**epresentations

Skill Learning Phase (G→B)



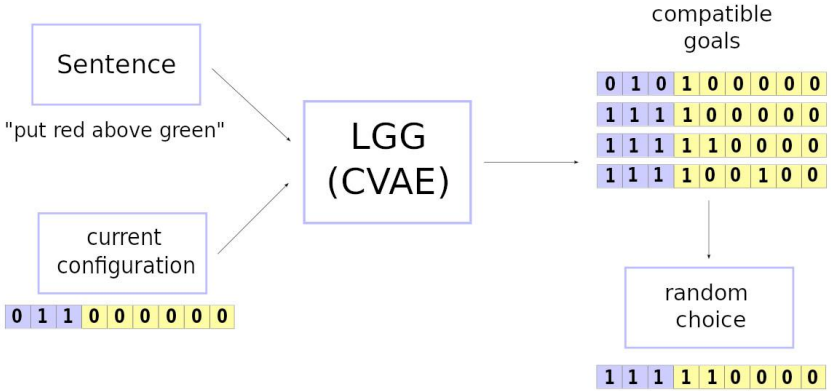


DECSTR uses 3 main components in the G→B phase:

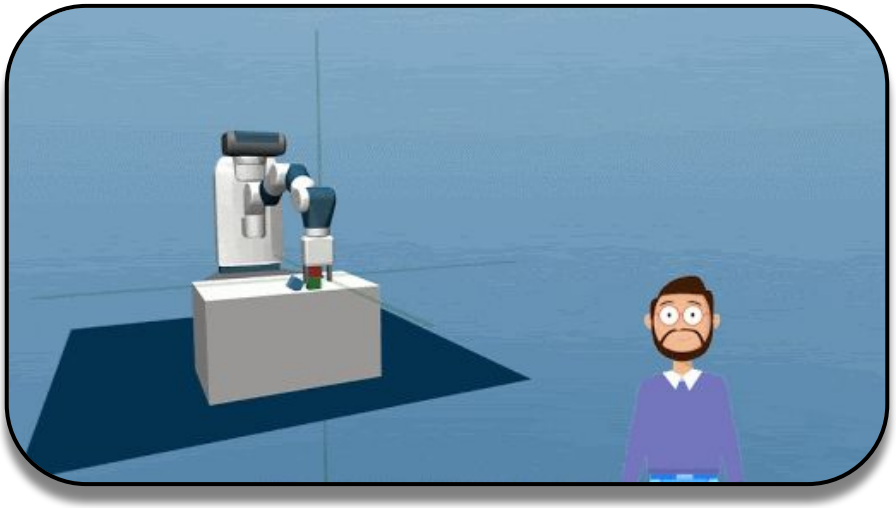
- Innate spatial predicate-based behavioral space (Mandler, 2012).
- Automatic Curriculum based on
  - aggregating of goals of similar difficulty.
  - selecting them according to the learning progress.
- Object-centered architecture that leverages Deep Sets (Zaheer et al., 2017).

Language Grounding Phase (L→G)

- A Social Partner (SP) describes one aspect of the transformations in object relations.

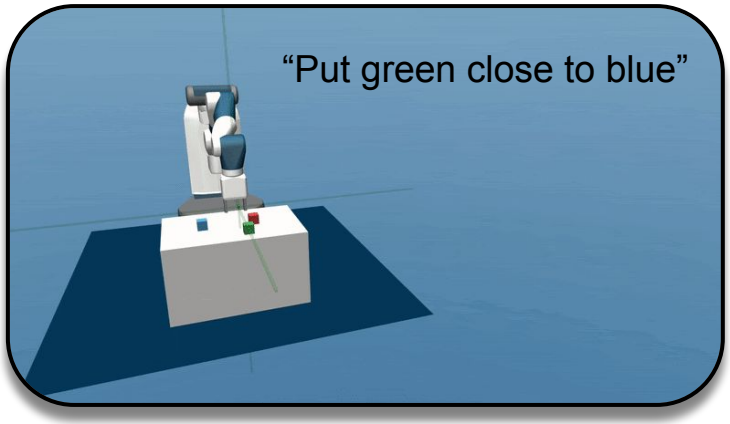


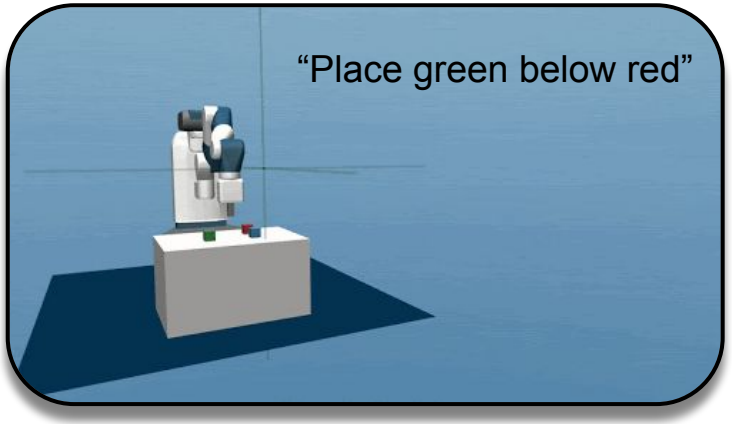
From an **initial configuration** and an **instruction**, it generates a **set of compatible goal configurations**.



- We create a dataset to train the Language-Conditioned Goal Generator (LGG).

Instruction Following Phase (L→G→B)

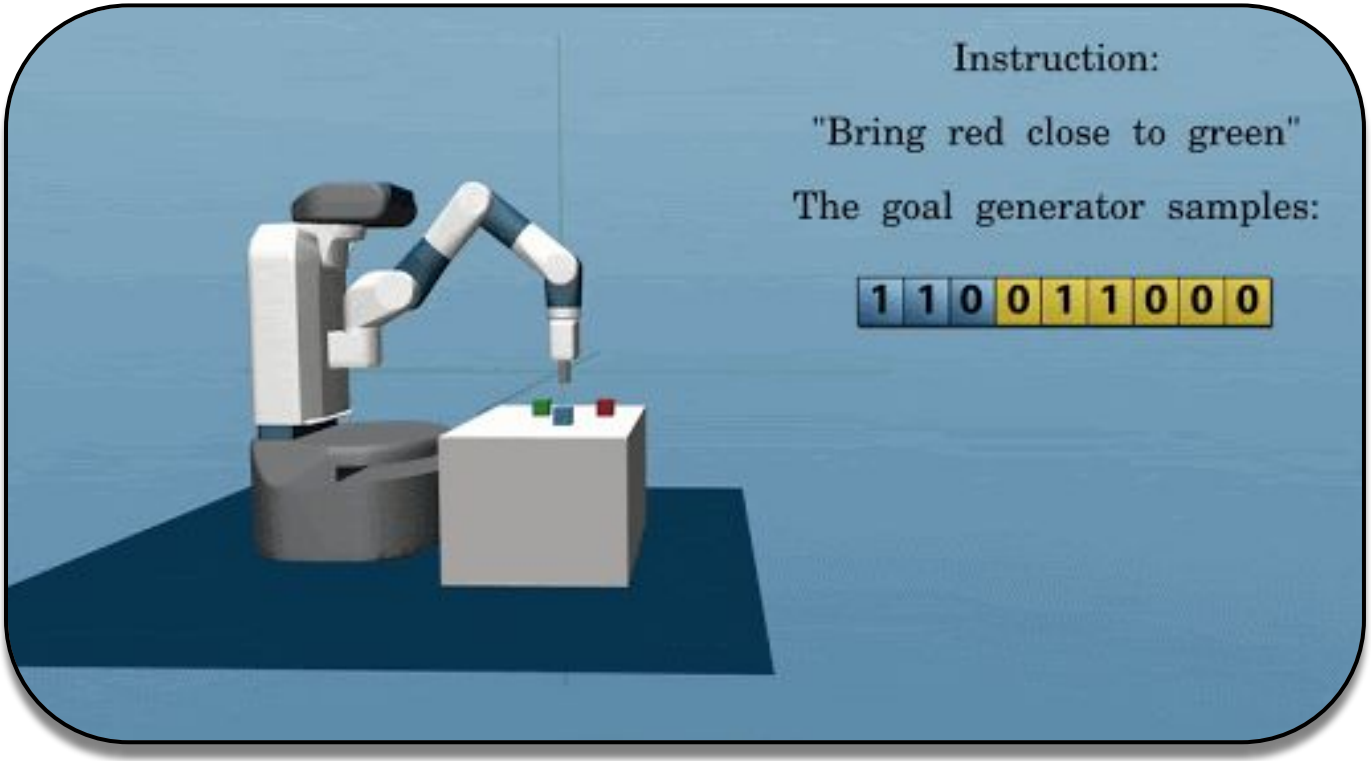




Now, DECSTR is ready to follow instructions !

- It receives **natural language-based instructions** from a SP.
- It maps them to **semantic goal configurations** that it then **pursues**.

Instruction Following Phase (L→G→B): Try again



Conclusion

The LGB architecture allows agents to learn alone or with a tutor through an intermediate semantic goal representation and a language-conditioned goal generator.

DECSTR is an instance of LGB that:

- masters predicate-based goals like preverbal infants (G⟶B),
- grounds language into its learnt behavior (L⟶G),
- pursues instructions with high behavioral diversity (L⟶G⟶B).

Come and have a chat at the virtual poster.

Arxiv ID: 2006.07185

Github repository: https://github.com/akakzia/decstr

Website: https://sites.google.com/view/decstr

Contact: ahmed.akakzia@isir.upmc.fr