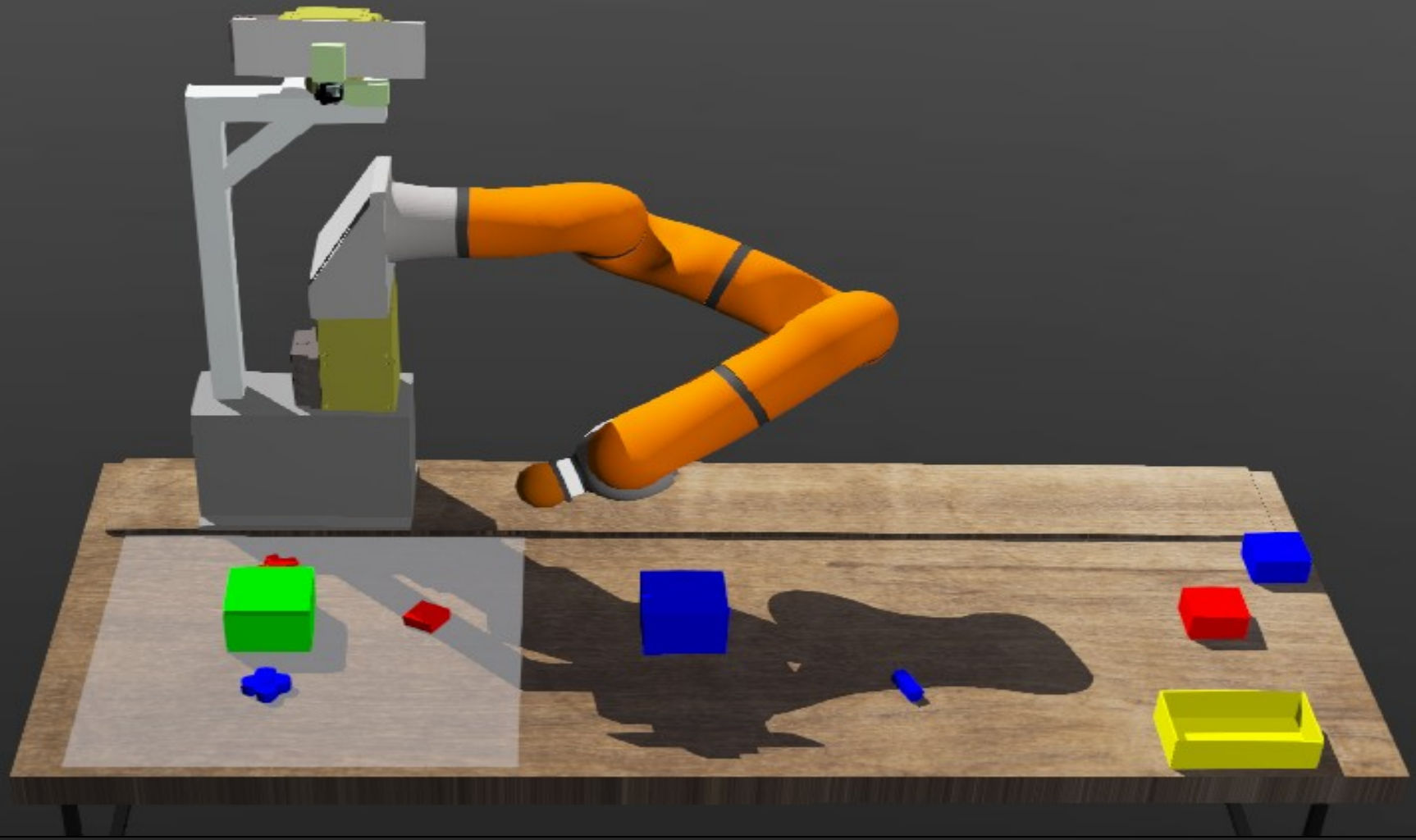


Action Grammar

FOR EMBODIED NEURAL AGENTS

Introduction

DFT-Research Group

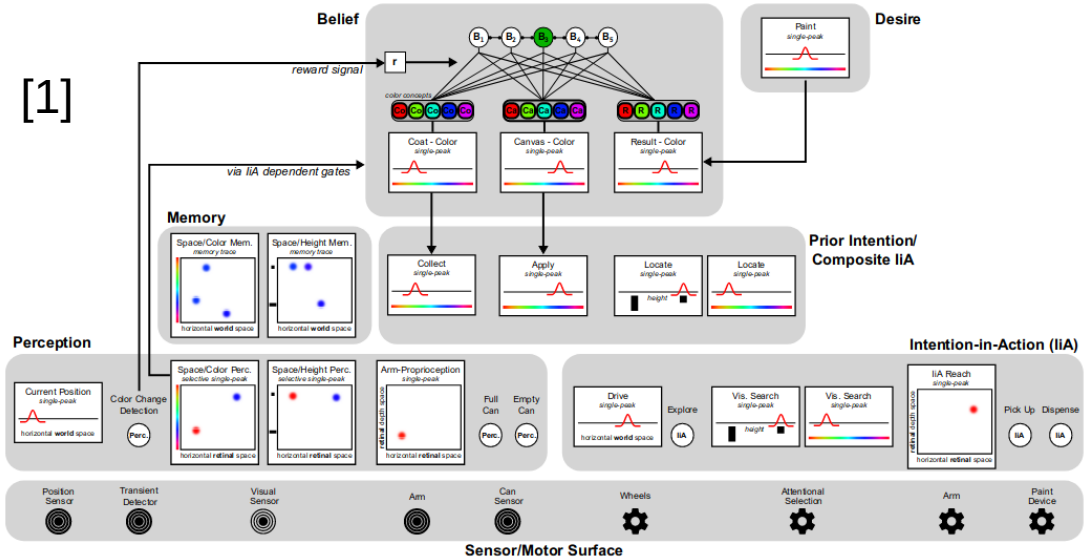
Intentional autonomous agents

(for example (Tekülve, 2022; Tekülve and Schöner, 2022; Tekülve et al., 2019)).

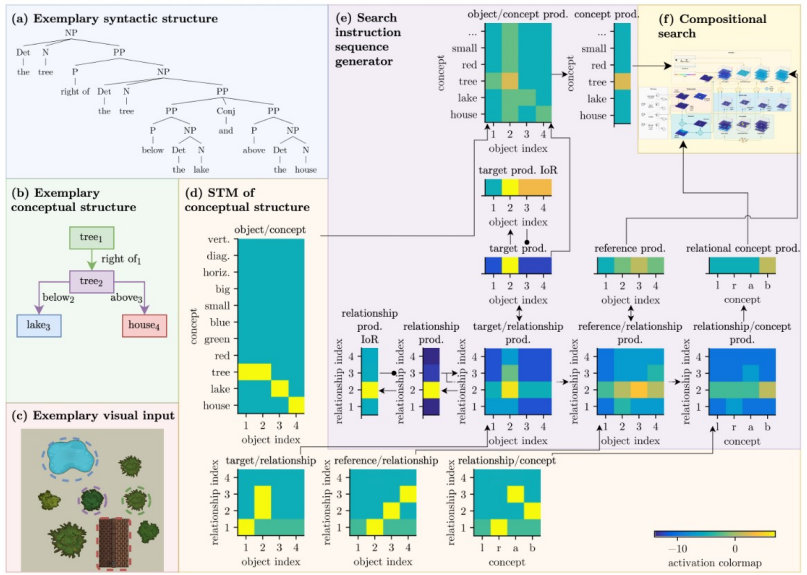
Perceptual grounding of language meaning

(for example (Richter et al., 2021; Sabinasz et al., 2020; Sabinasz and Schöner, 2022)).

[1]

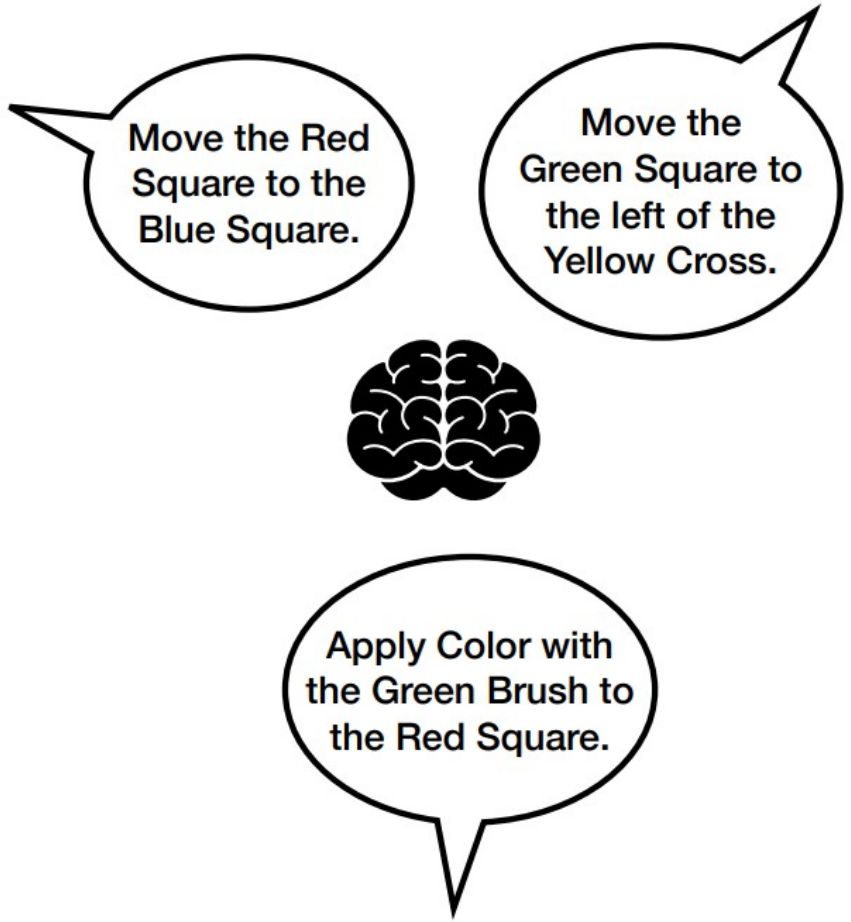


[4]



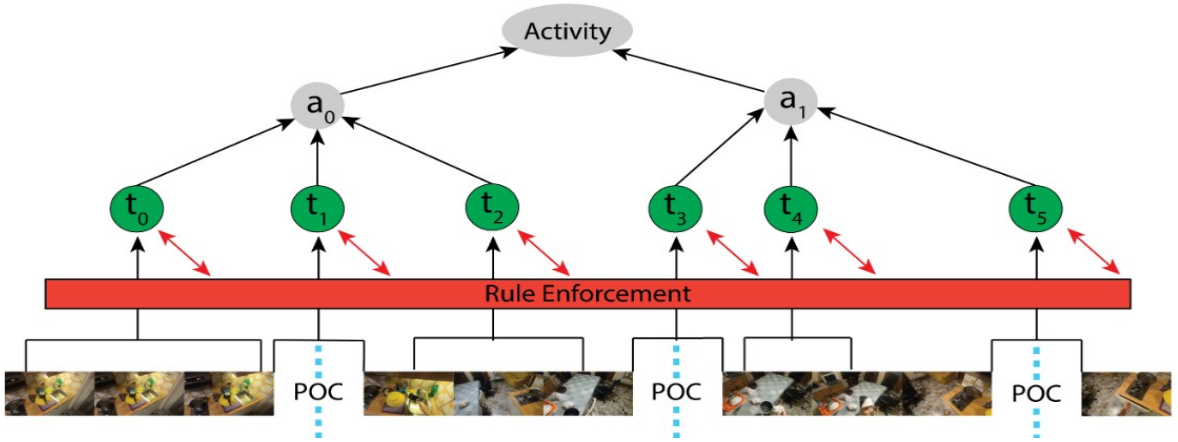
Action Grammar for Embodied Neural Agents

- Like language, human actions seem to possess a high degree of flexibility and seem to exhibit a similar compositional and hierarchical structure.
- As analogous to linguistic expressions, the sequential arrangement of actions reveals a structure that seems to be in line with the compositional and hierarchical organization of language.

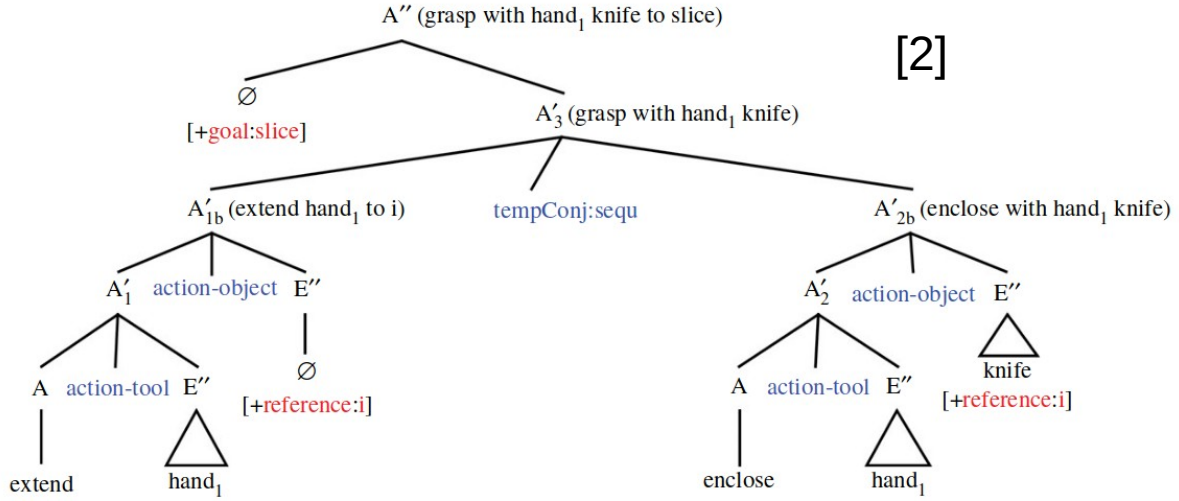


Minimalist Grammar of Action

- Computer vision: Action Parsing based on points of contact.
- Production rules about how components can be merged to produce Action Phrases:
  - Action primitives detected based on points of contact
  - Tools, objects and primitives combined into a part of parsing tree
  - Sequential buildup of parsing tree to overall activity as action unfolds

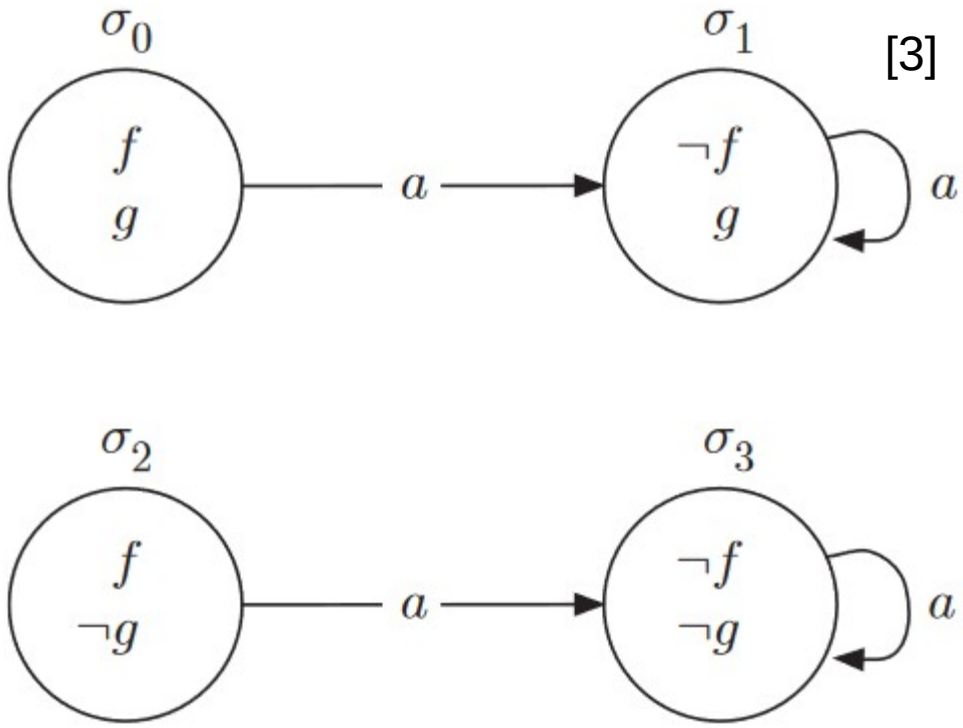


[2]



Action Description Languages

- Classic STRIPS style planning:
  - Utilizes propositional logic to create a formal language.
  - Statements about cause and effect can be used to reason and plan inside the domain of the language.
- Predicates are arbitrary disembodied atomic symbols
  - Leads problems such as the ramification problem



[3]

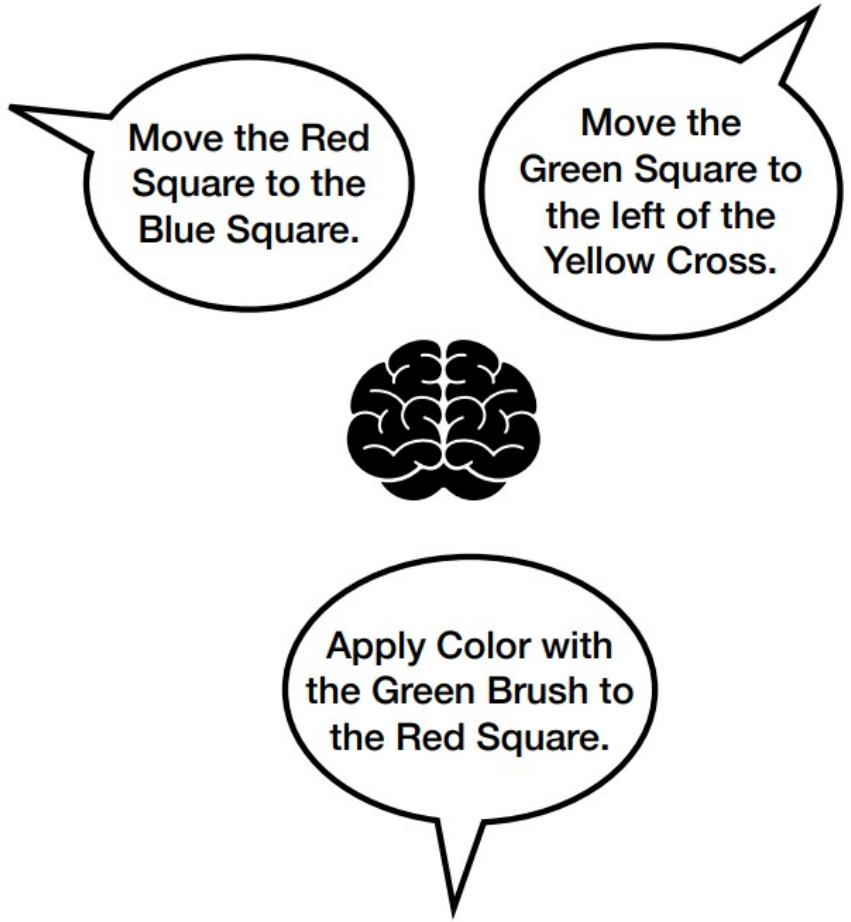
- Could an embodied action language be a candidate to explain reasoning and planning behavior of humans?

$\mathcal{F} = \{f, g\}\ \mathcal{A} = \{a\}$

a **causes** $\neg f$ **if** f

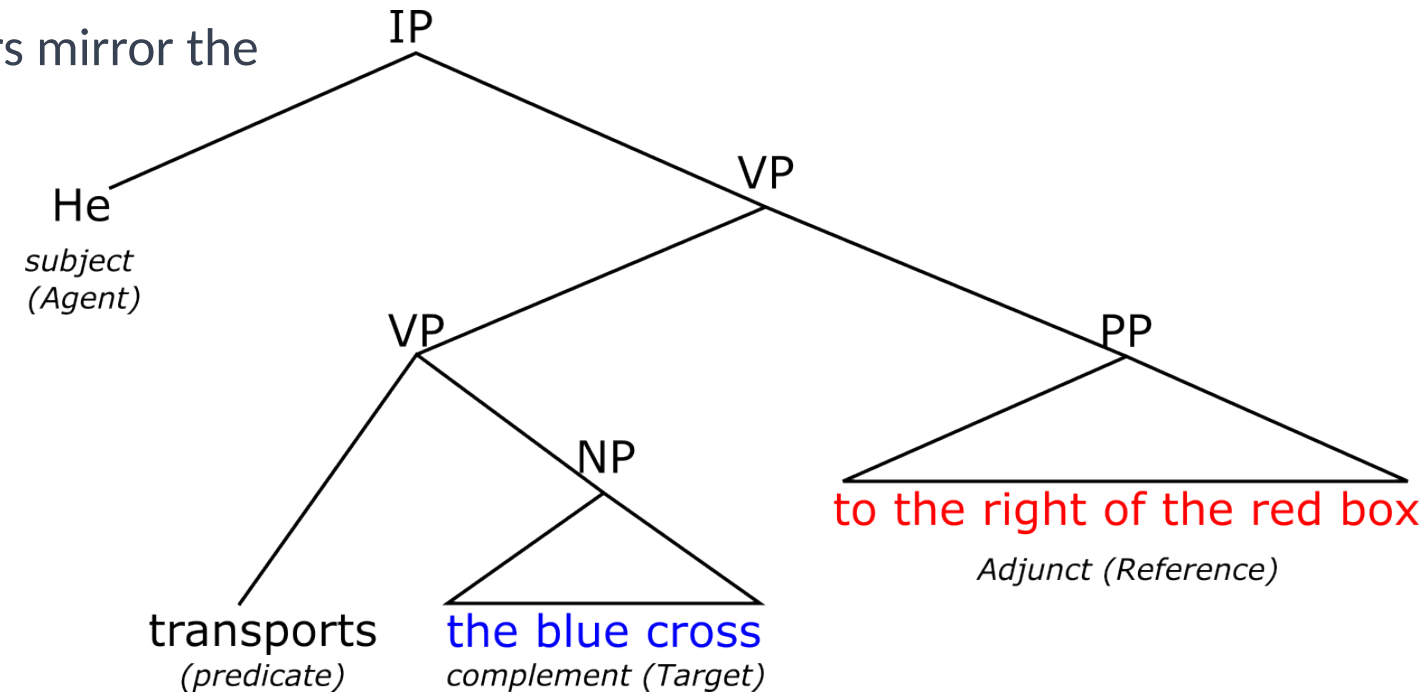
Action Grammar for Embodied Neural Agents

- Hypothesis: There exists a common cognitive structure underlying structured representations in both language and action. The Conceptual Structure.
- This hypothesis is suggested by resemblances in their compositional and hierarchical structure.
- Goal: Extend the intentional states of action to the conceptual structure, to allow flexible action composition and planning.



Structure of Action Phrases

- Action verbs derive their complete meaning through the inclusion of specific arguments.
- The valency of verbs varies, encompassing monovalent, divalent, trivalent forms, and more.
- Complements, Adjuncts, and other specifiers mirror the role of semantic roles of verbs.
- Argument Structure:
  - Verb (head)
  - Semantic Roles (Complements)
    - Target, Reference, Agent, …
  - Optional Modifiers (Adjuncts)



Minimal imperative Action Language

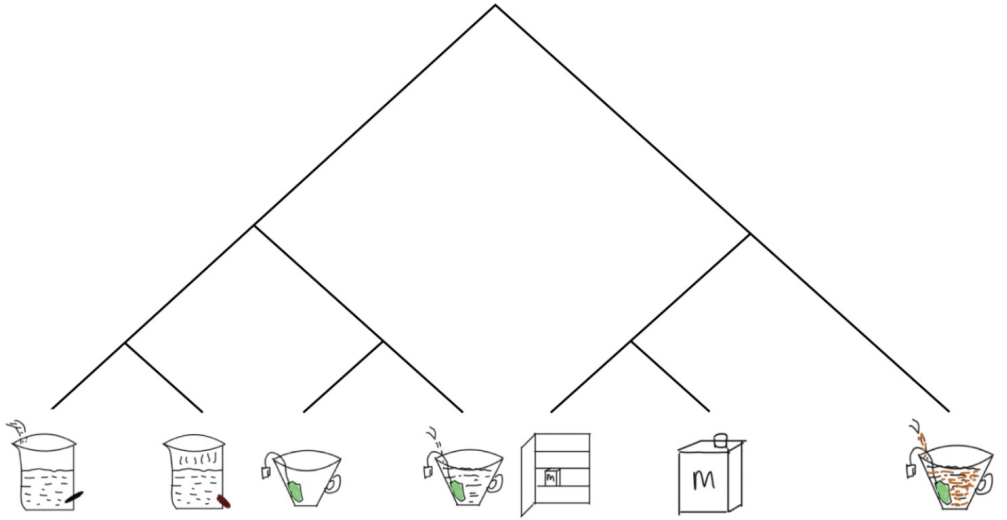
- Action Phrases (APs) with fixed set of Action Roles:

ACTION	TARGET	REFERENCE	RELATION
Verb	„what“	„where“	„modifier“

- "**Transport** [Action] **the Red Cross** [Target] **to the left of** [Relation] **the Green Box** [Reference]“
- “**Apply Color** [Action] with **the Red Brush** [Target] to **the Blue Flower** [Reference].”

Hierarchical Action Alphabet

- Higher level actions can be decomposed into sub-actions geared towards specific sub-goals.
- A sequential arrangement of lower level sub-actions contributes to the accomplishment of the overarching task.
- The Verb in the higher level action phrase serves as a label for a set of lower level action sequences.
- Equivalent sequences that fulfill the same overarching goal can be grouped under the same label.



Hierarchical Action Alphabet

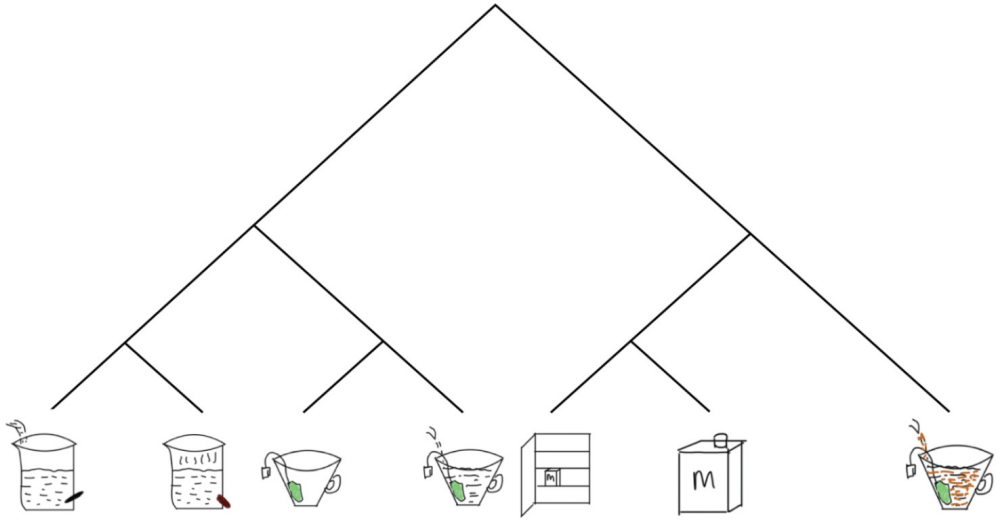
- Example: Making Tea

A = {make_tea, boil_water, put_tea_in_cup, fill_cup, fill_kettle, start_kettle}

S_boil_water -> {(fill_kettle, start_kettle)}

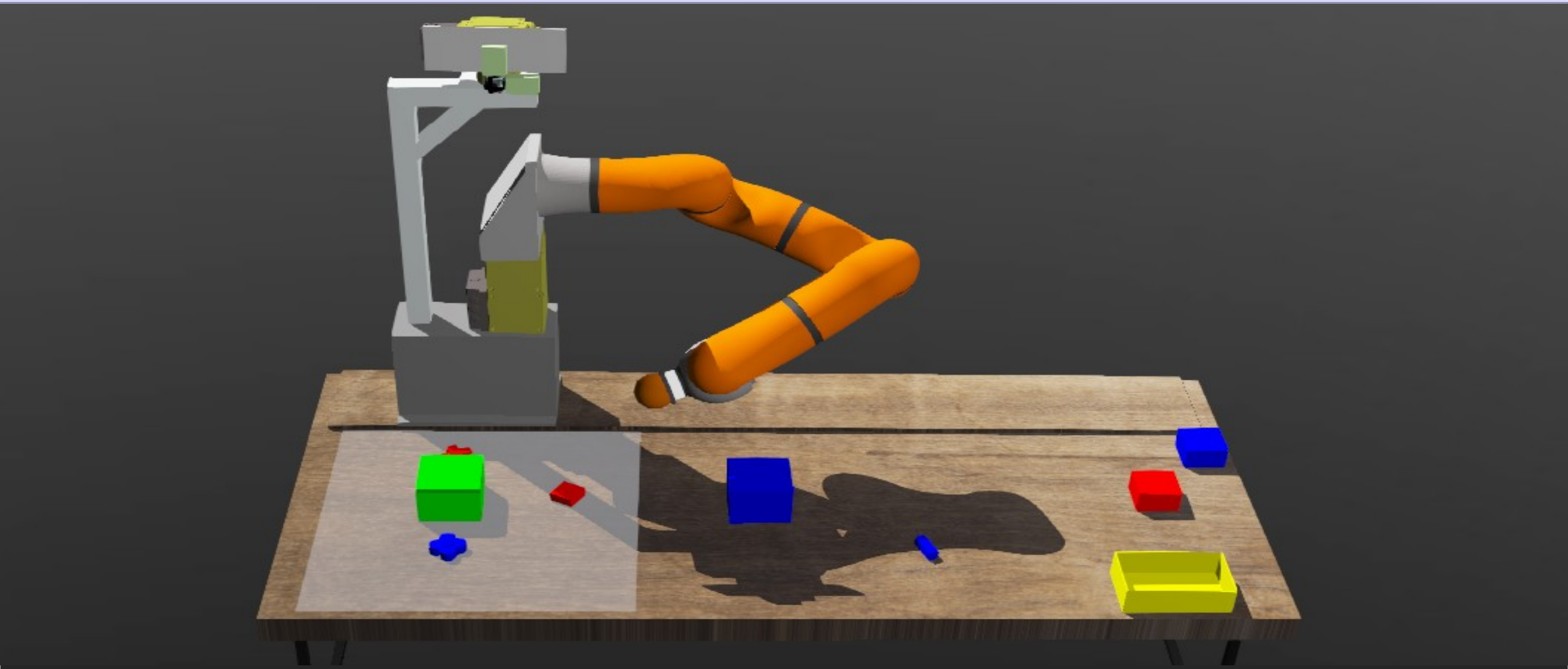
S_make_tea -> {(put_tea_in_cup, boil_water, fill_cup), (boil_water, put_tea_in_cup, fill_cup), (fill_kettle, put_tea_in_cup, start_kettle, fill_cup), ...}

$S\_make\_tea \rightarrow (boil\_water * fill\_cup) \otimes put\_tea\_in\_cup$



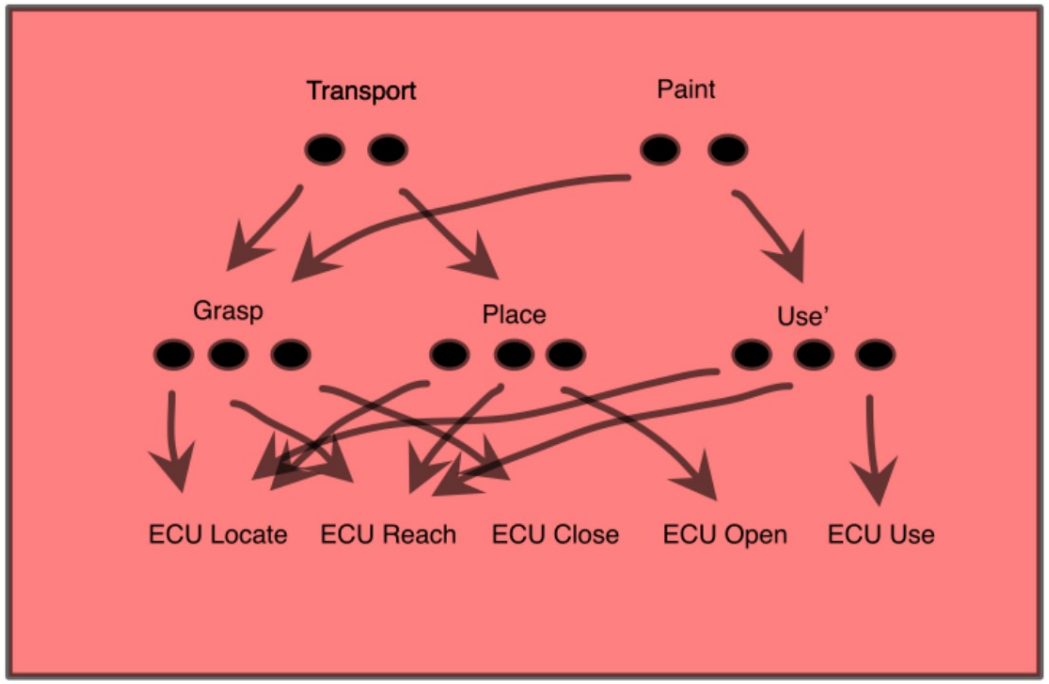
[5]

Grounding Hierarchical Action Grammar



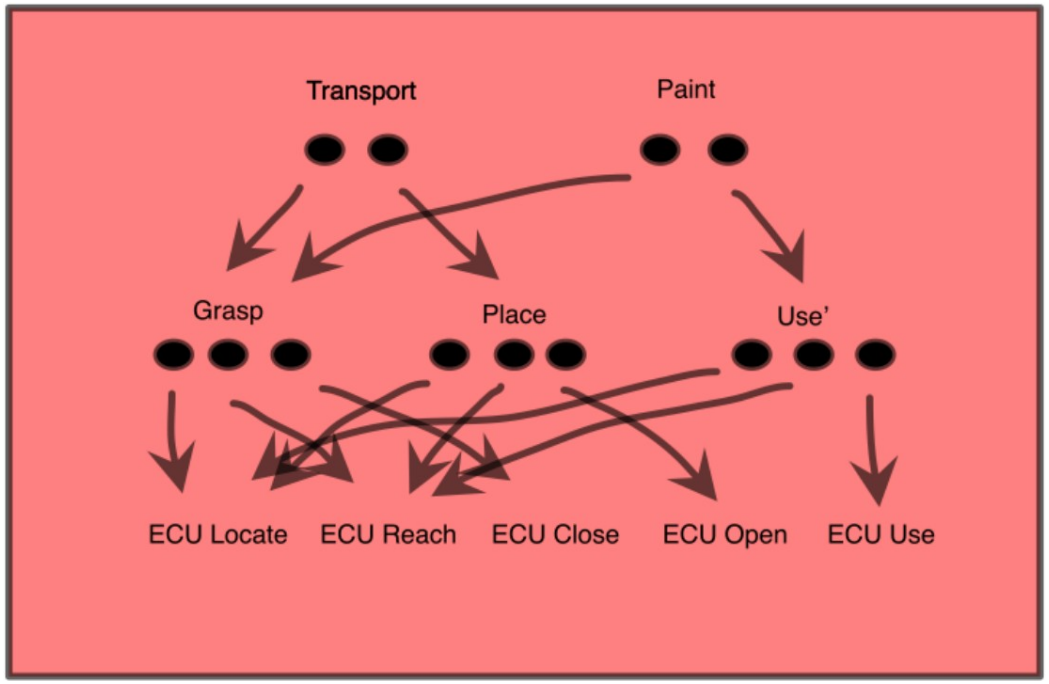
Hierarchical Action Alphabet

- We take elements of our action alphabet to be already learned action concepts that can be executed without additional planning.
- Hierarchical action concepts that point to lower level motor schemas.



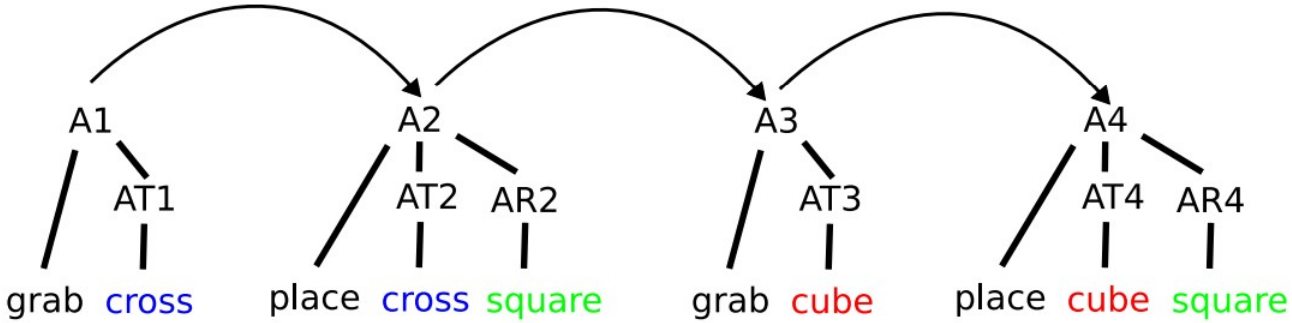
Hierarchical Action Alphabet

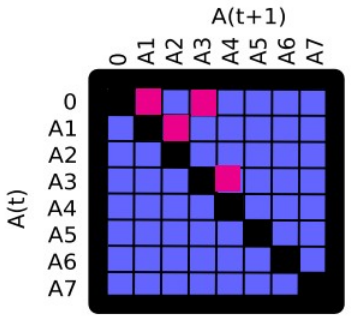
- We take elements of our action alphabet to be already learned action concepts that can be executed without additional planning.
- Hierarchical action concepts that point to lower level motor schemas.
- Action Alphabet: {locate, reach, open, close, use, grasp, place, transport, paint}
  - Grasp -> (reach, close)
  - Place -> (reach, open)
  - Transport -> (grasp, place) -> (reach, close, reach, open)

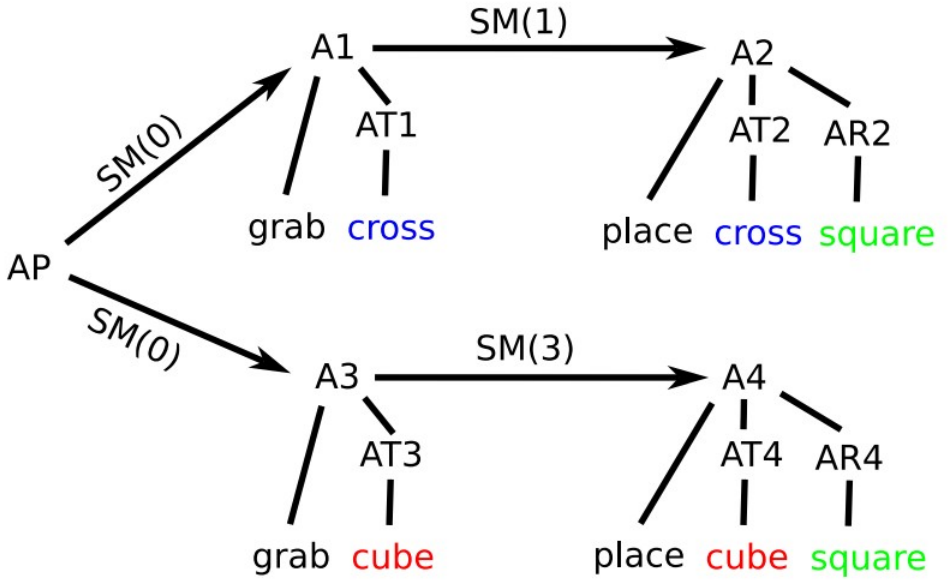


Action Plan

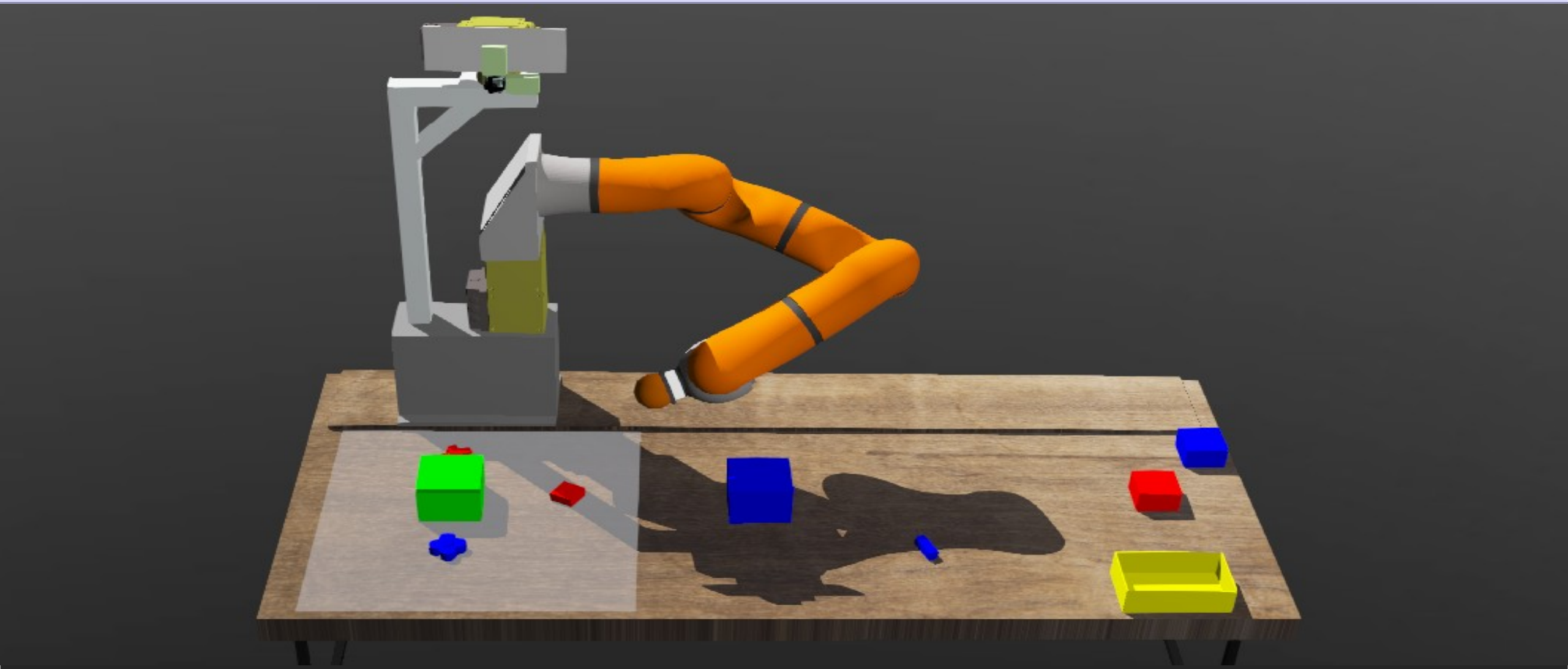
- Action concepts point to fixed grounded sequences of lower level actions.
- Humans show flexibility in action execution.
  - Opportunistic for independent actions
  - Sequential for order dependent actions
  - -> partial ordering
- Successormap to represent dependency relations in given action plan.



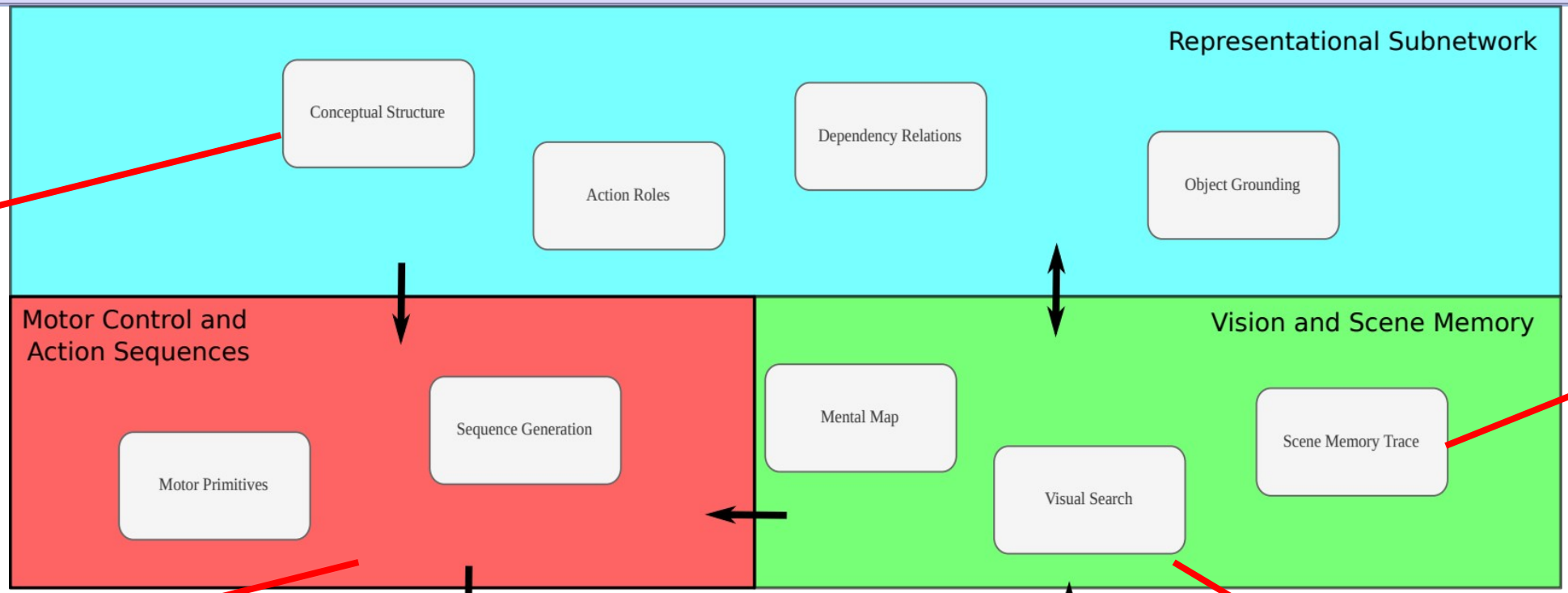




Preliminary Results



Integrative Architecture

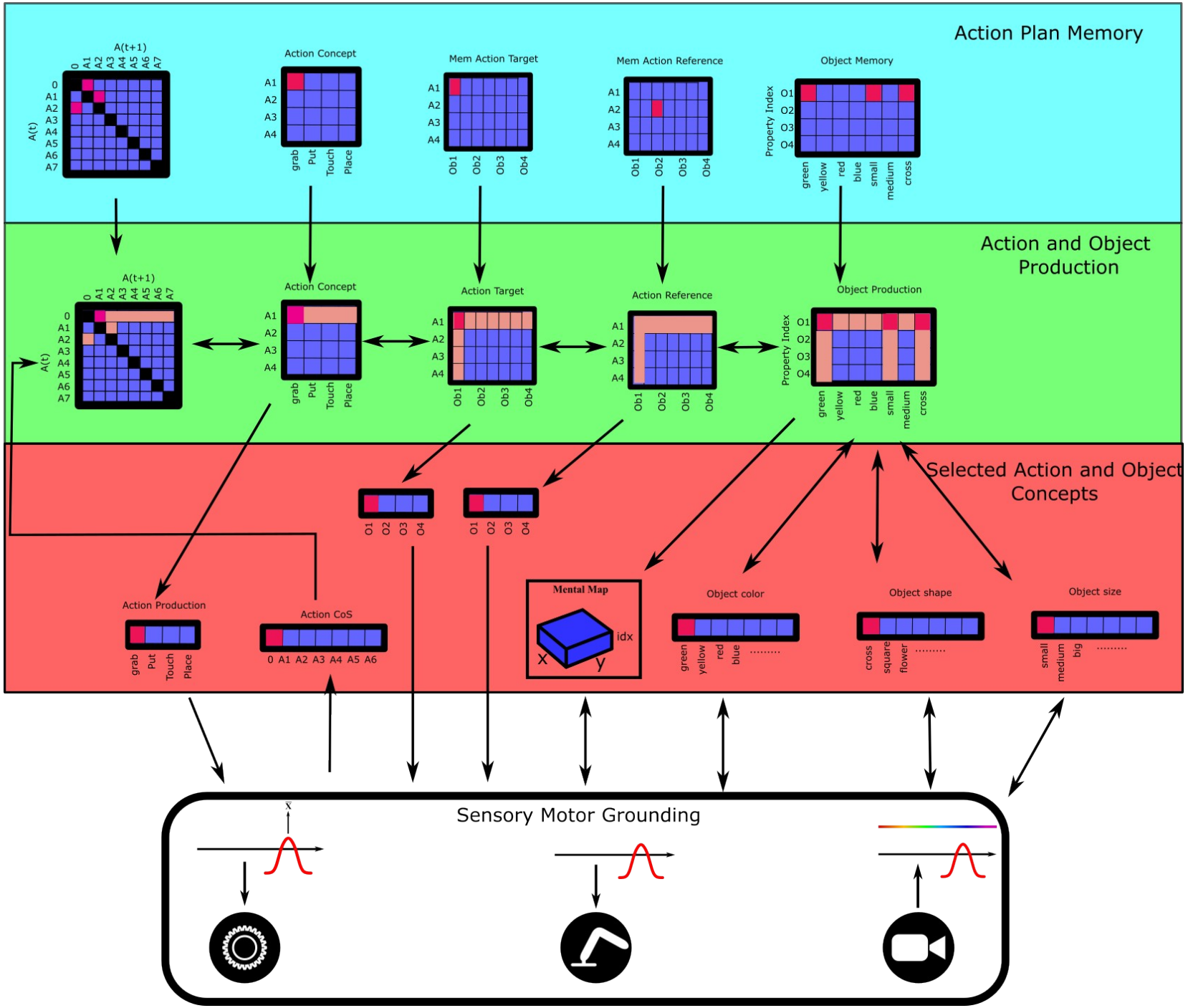


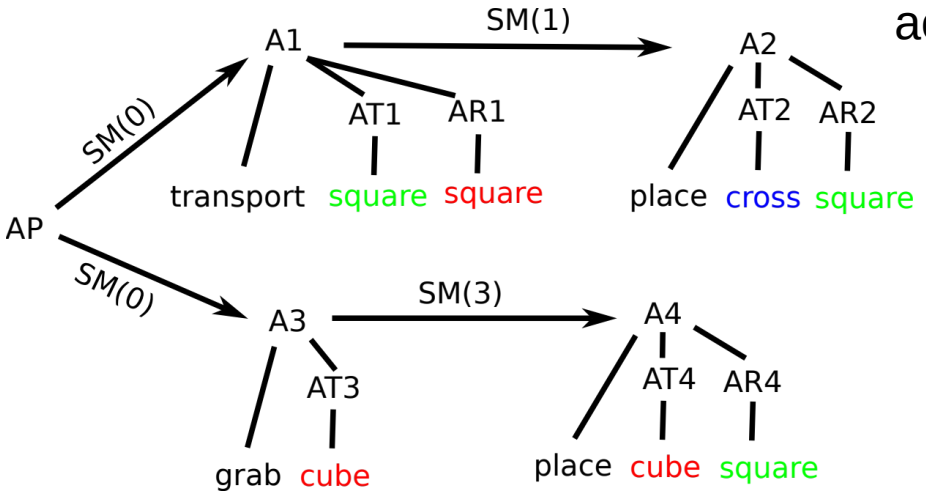
Sabinasz et al. (2022)

Ziebner et al.(2010)

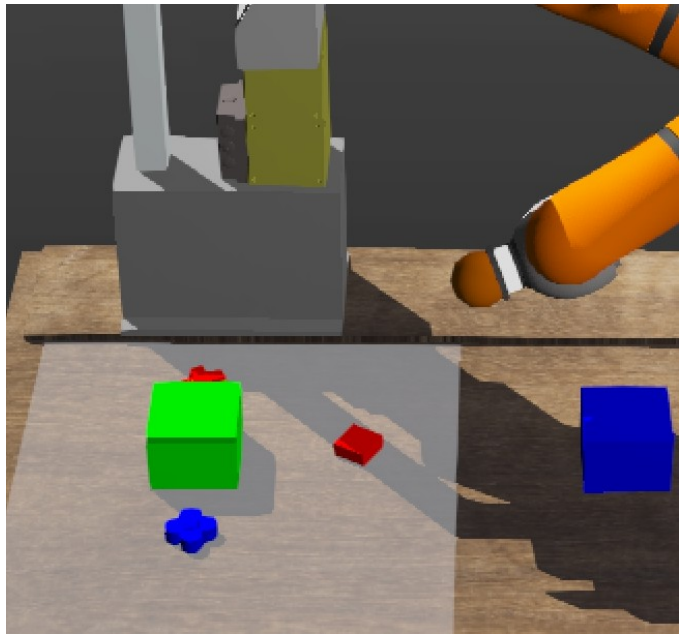
Tekuelve (2022)
Sandamirskaya (2011)

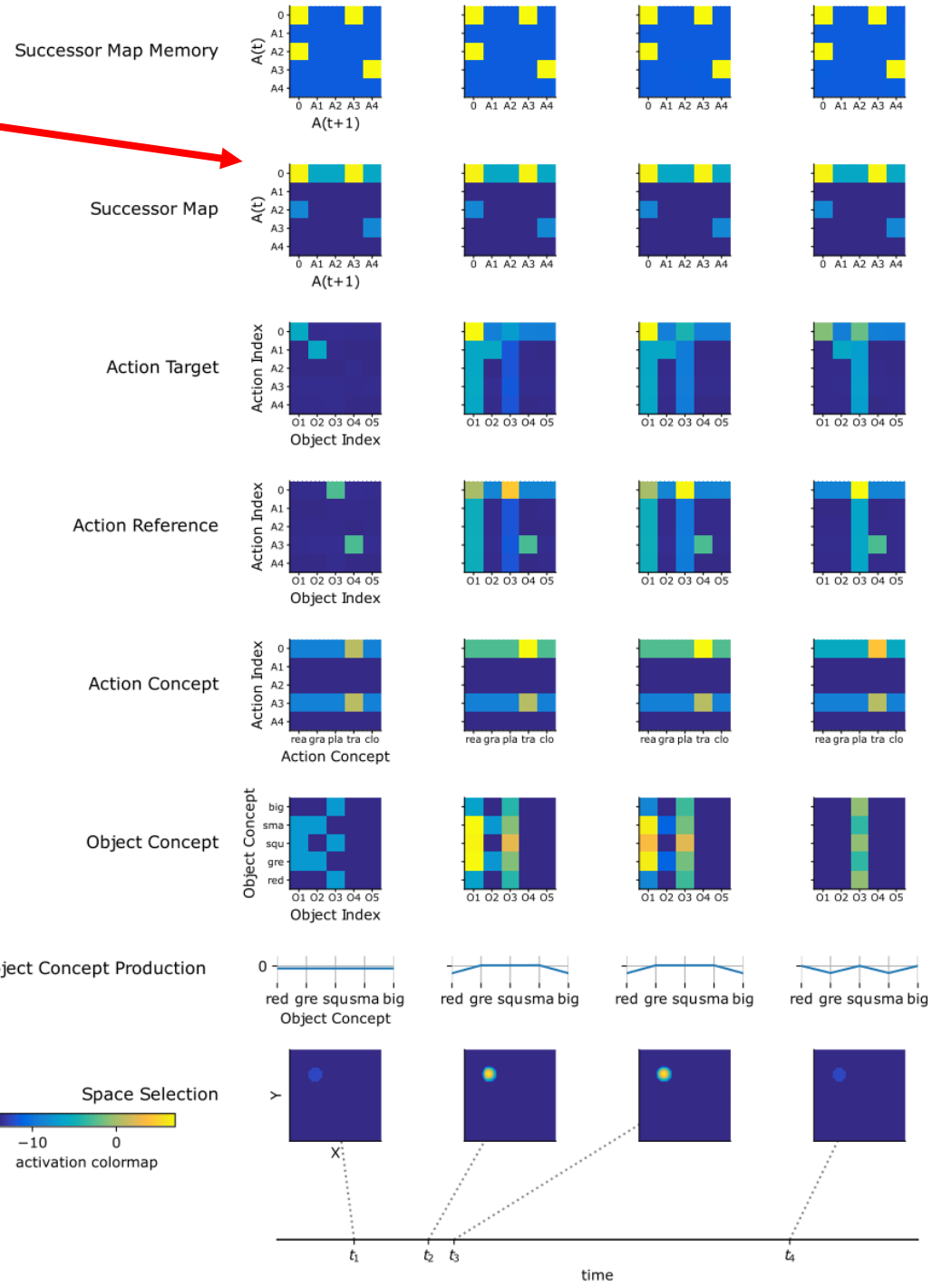
Richter et al. (2012)
Grieben et al. (2020)

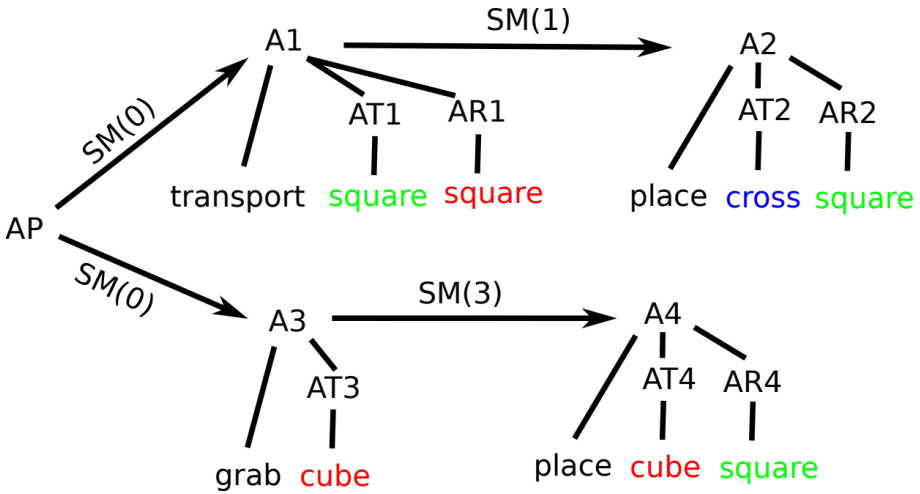

Action Plan Memory
A(t+1)
A(t)
Action Concept
grab
Put
Touch
Place
Mem Action Target
Mem Action Reference
Ob1
Ob2
Ob3
Ob4
Object Memory
Property Index
O1
O2
O3
O4
green
yellow
red
blue
small
medium
cross
Action and Object Production
Action Target
Action Reference
Object Production
Selected Action and Object Concepts
Action Production
Action CoS
0 A1 A2 A3 A4 A5 A6
Mental Map
idx
x
y
Object color
Object shape
Object size
square
flower
big
Sensory Motor Grounding

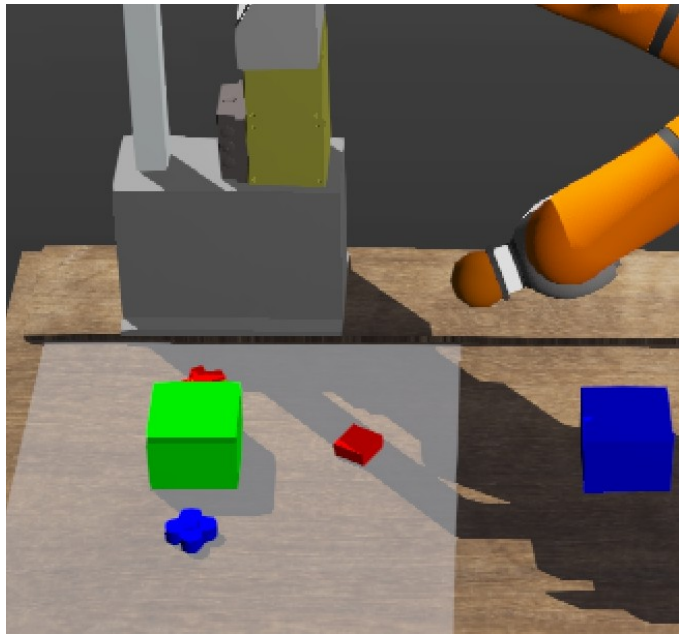


SM preshapes possible action phrases



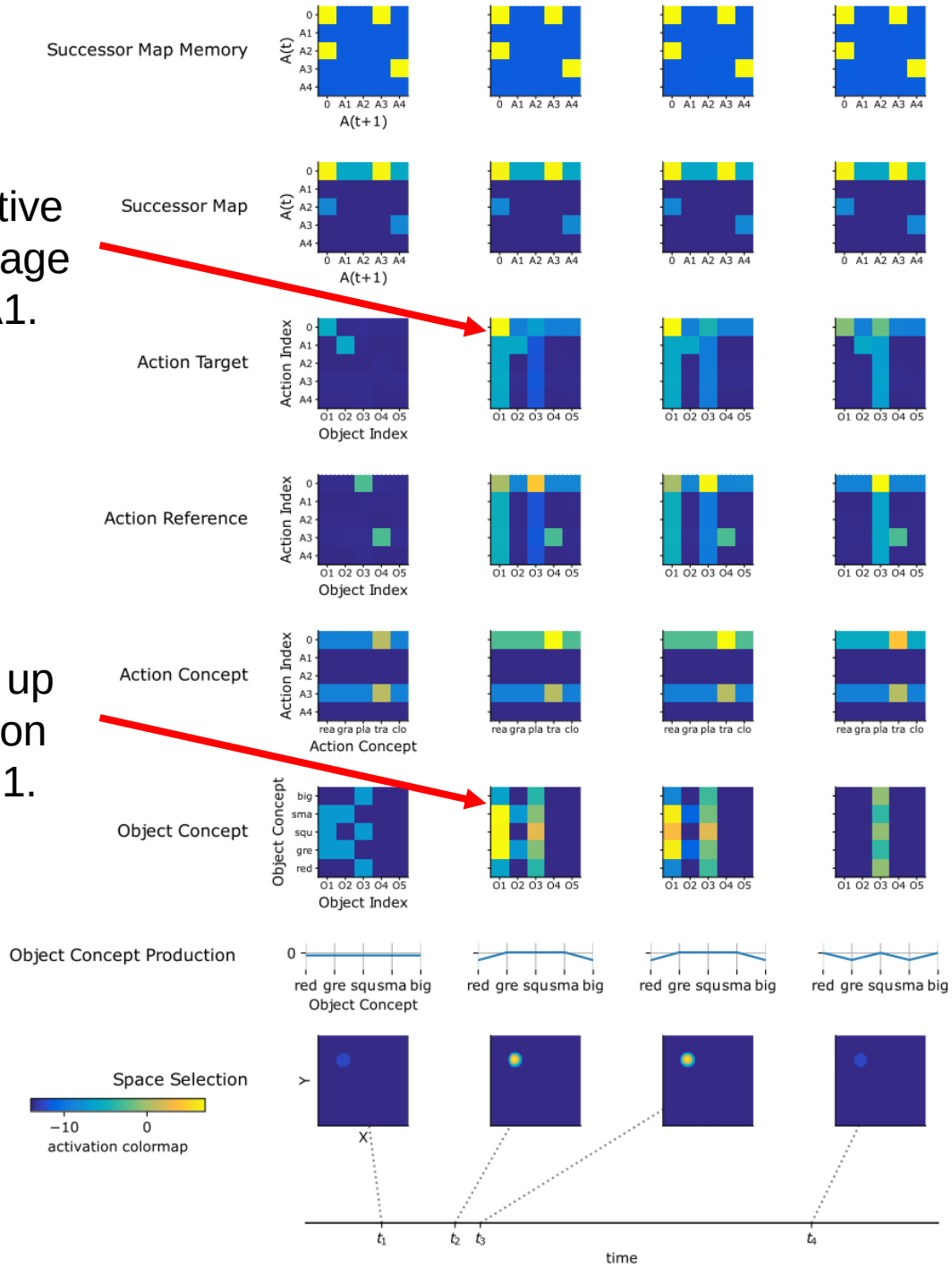


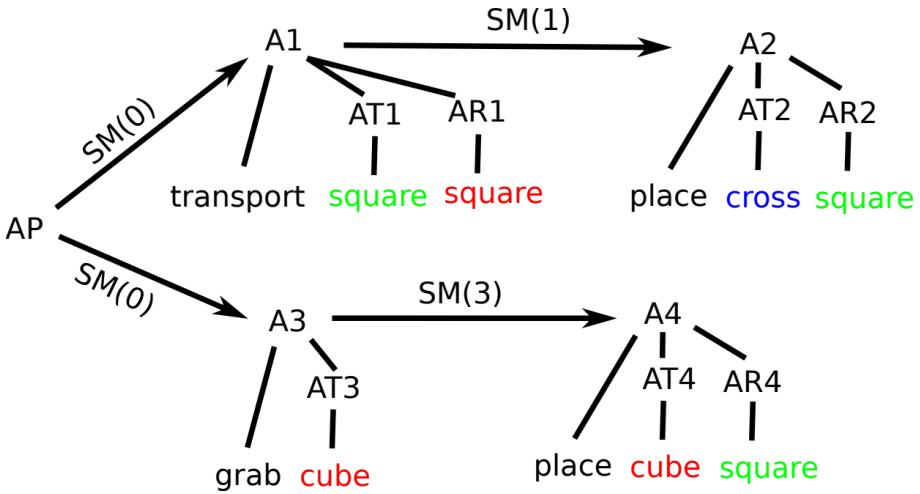


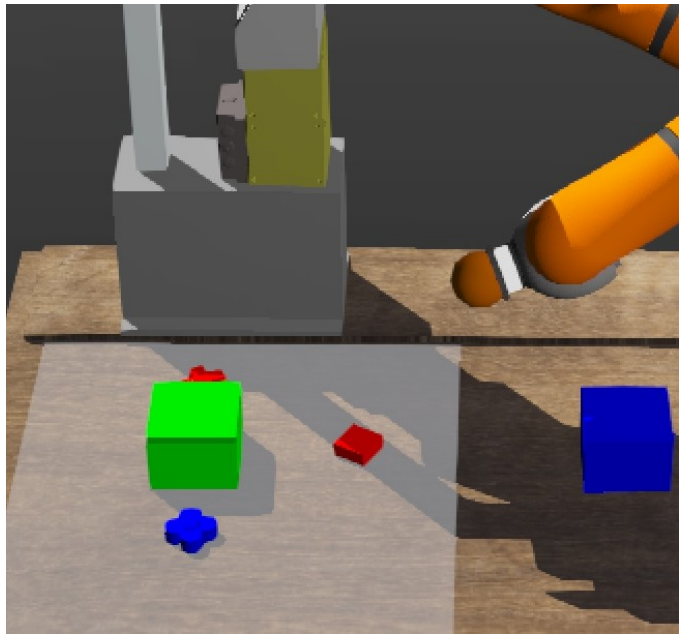


Selective advantage for A1.

Bottom up selection of Obj1.

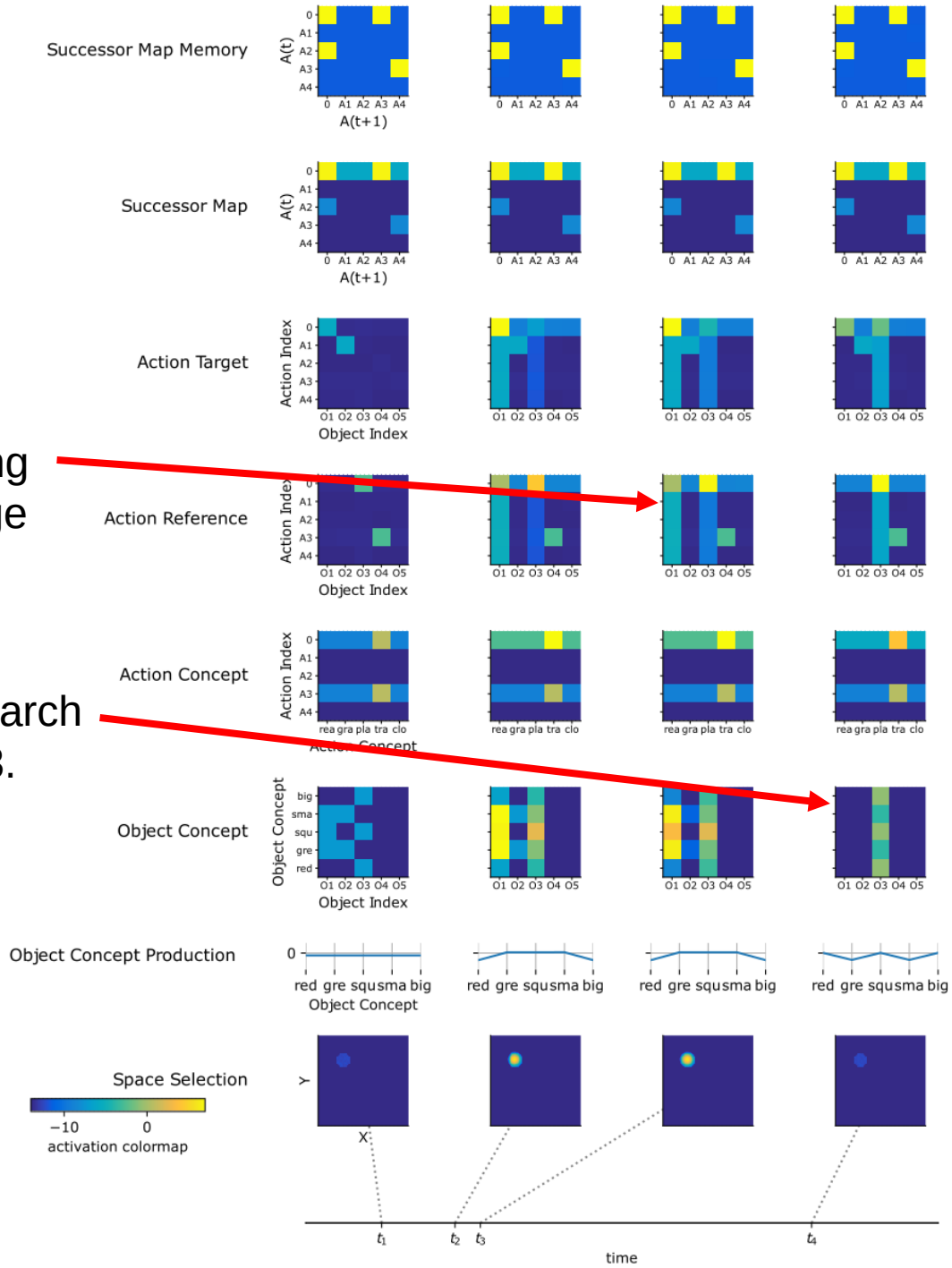






After grounding advantage for O3.

Visual Search for O3.



Conclusion

- Minimal imperative action language for embodied agent robot
- Cognitive structure for representations of structured action phrases and dependency relations in action plans
- Opportunistic and flexible grounding of action plans

- Future questions:
  - How can action concepts be learned?
  - Action language for reasoning and Planning possible?

Literature

- J. Tekülve, “A Neural Process Model of Intentionality Implemented on an Autonomous Robot”, (03.2021)
- K. Pastra, “The minimalist grammar of action”, (2012), Philosophical Transactions of the Royal Society B: Biological Sciences
- Michael Gelfond, “Knowledge Representation, reasoning and the design of intelligent agents: the answer set programming approach”, (2014), Cambridge University Press.
- D. Sabinasz, “A Neural Dynamic model perceptually grounds nested noun phrases”, (2022), Proceedings of the 44th annual conference of the cognitive science society
- C. Coopmans et al., “Hierarchical structure in action and language: a formal comparison”, (2023), Psychological Review