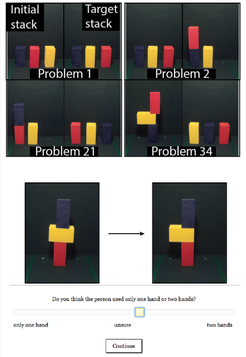
Physical problem solving: Joint planning with symbolic, geometric, and dynamic constraints

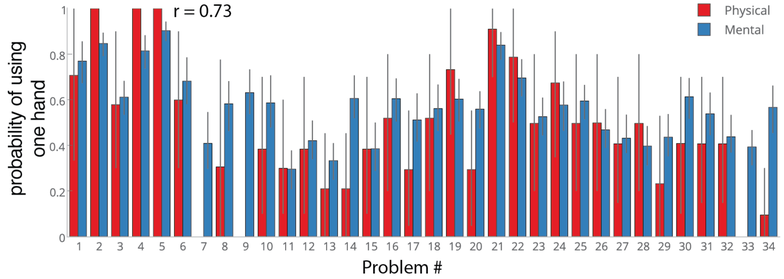
Ilker Yildirim*, Tobias Gerstenberg*, Basil Saeed, Marc Toussaint, Joshua B. Tenenbaum
Department of Brain & Cognitive Sciences, MIT

Introduction

- Physical problem solving is a core aspect of intelligence and ubiquitous in everyday cognition: from young children stacking cups, to adults moving furniture to redesign a room. Human-like flexible problem-solving heavily draws upon common-sense reasoning and model building.
- We introduce a new paradigm, the stack re-configuration problems, to study human physical problem solving in complex sequential object manipulation tasks.
- Computational requirements: (1) An abstract high-level symbolic plan (e.g., place the red block on top of the blue block, etc.), (2) executing such plans taking into account the geometric (e.g., positions, shapes, etc.) and (3) dynamic constraints (e.g., physical stability, weight of objects, etc.).
- We draw inspiration from two separate fields to satisfy these requirements: (1) research on mental simulation, and (2) optimal control theory.

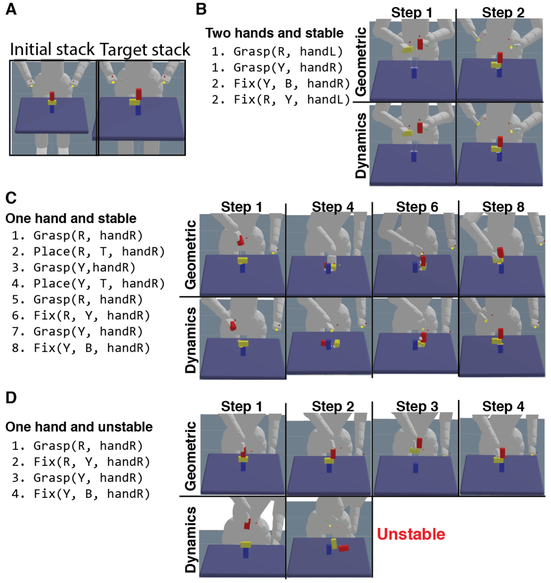
Stack re-configuration problems: How many hands needed, one or two?





Computational model: Integrating symbolic, geometric, and dynamic constraints

- Symbolic plans are a sequence of abstract motion primitives (`Grasp`, `Fix`, `Place`). We use Monte Carlo Tree Search to satisfy symbolic constraints.
- Geometric constraints are used to find hand and object trajectories under a given symbolic plan. We use the KOMO solver to satisfy these constraints.
- Dynamic constraints are used to test the solutions for physical stability. We use a physics engine to satisfy these constraints.

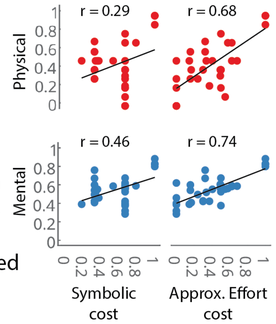


Results: Full and lesioned model

- Symbolic cost, s_{ij}: Each solution, j, for each problem, i, is assigned a cost as the number of steps in its symbolic plan.

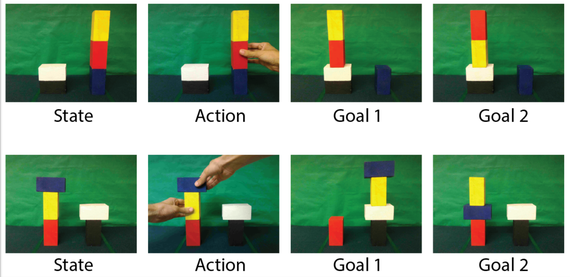
$$Pr_s(OneHand|Problem_i) = exp(-\beta * \frac{1}{K} \sum_{j \in one\_hand\_solutions}^{K} s_{ij}) / (exp(-\beta * \frac{1}{K} \sum_{j \in one\_hand\_solutions}^{K} s_{ij}) + exp(-\beta * \frac{1}{L} \sum_{j \in two\_hands\_solutions}^{L} s_{ij}))$$

- Approximate effort cost, f_{ij}: Each solution, j, for each problem, i, is assigned a cost as the total amount of kinetic energy required to carry out the solution in simulation. Moreover, physically unstable solutions are eliminated.

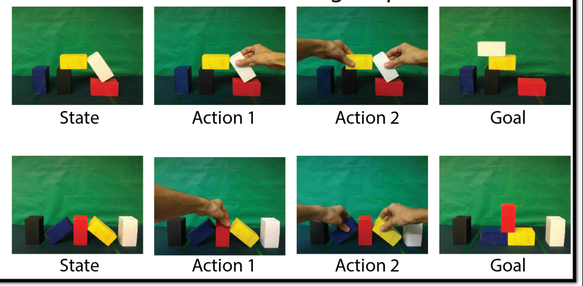


Future directions: Goal & action inference

Which goal does the agent pursue?



Which action did the agent perform?



Conclusions

- Stack-reconfiguration problems allow us to quantitatively study how people plan and execute non-trivial sequential object manipulation tasks, and how they reason about others' goals and actions in a "theory of mind" context.
- A model combining symbolic, geometric, and dynamic constraints accounts well for the data. People's judgments cannot be predicted by merely considering symbolic costs.
- A substantial challenge for the current dominant approach in AI, i.e., deep learning, is to engineer data-efficient and flexible physical problem solvers.

Acknowledgements

This work was supported by the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216 and by an ONR grant N00014-13-1-0333.