SEE, FEEL, ACT: HIERARCHICAL LEARNING FOR COMPLEX MANIPULATION SKILLS WITH MULTISENSORY FUSION

RUPRECHT-KARLS-UNIVERSITÄT HEIDELBERG

FAKULTÄT FÜR MATHEMATIK UND INFORMATIK

INSTITUT FÜR INFORMATIK

SEMINAR: AI IN GAMES

DOCENT: PROF. DR. U. KÖTHE

SPEAKER: KERSTIN KÜSTERS

CONTENTS

- Introduction
- Simulation
- Challenges
- How to play Jenga as a robot
- Let‘s play Jenga
- Conclusion

Fig. 1: A Jenga playing robot.

INTRODUCTION

- Published in the Science Robotics last January
- Six researchers: N. Fazeli, M. Oller, J. Wu, Z. Wu, J.B. Tenenbaum, A. Rodriguez

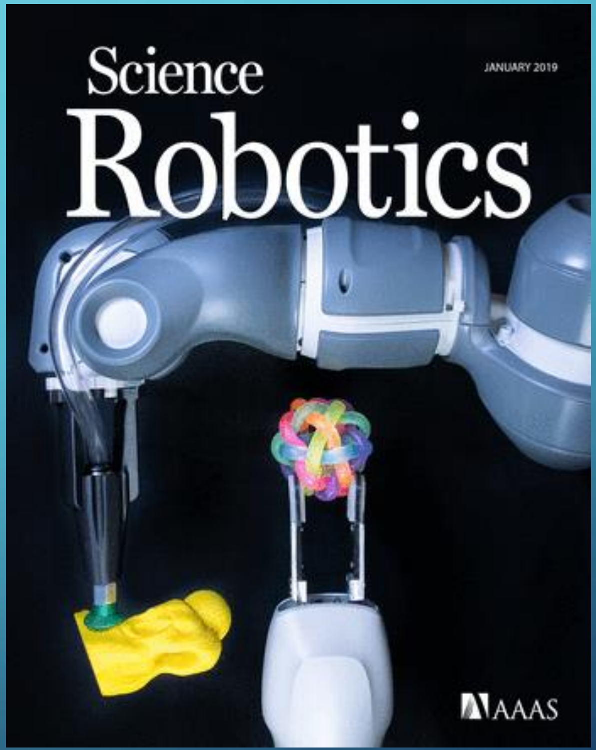


Fig. 2: Vol. 4 Issue 26 of the Science Robotics.

INTRODUCTION

- Humans combine different senses to learn
    - One of the most effective learning techniques
    - Visual data is often the primary input used in AI
    - Current learning algorithms struggle with…
        - …active perception…
        - …and hybrid behavior
    - And don’t exploit physics enough
- Why Jenga?
- Jenga rules and robot rules

SIMULATION

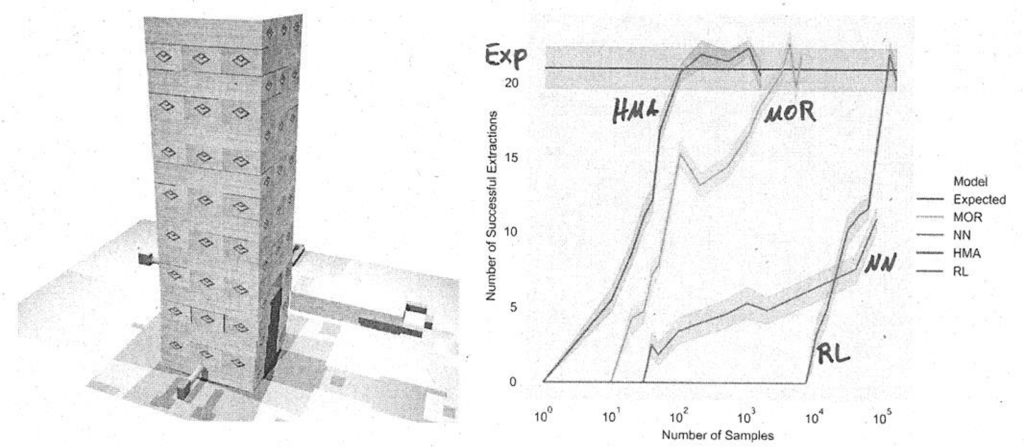


Fig. 3: Baseline comparison.

- Baseline comparison
  - Hierarchical Model Abstraction (HMA)
  - Feed-forward Neural Network (NN)
  - Mixture of Regressions (MOR)
  - Proximal policy optimization (PPO) implementation of Reinforcement Learning (RL)

CHALLENGES

- Feel if a block is movable or immovable
- Necessary force and angle to move a block without damaging the tower
- Visual system limitations
- Many contact-rich manipulation skills are difficult to automate for large-scale data collection
- Sim-to-real transfer of policies
- An effective integration of tactile information and the persistent visual stream

HOW TO PLAY JENGA AS A ROBOT

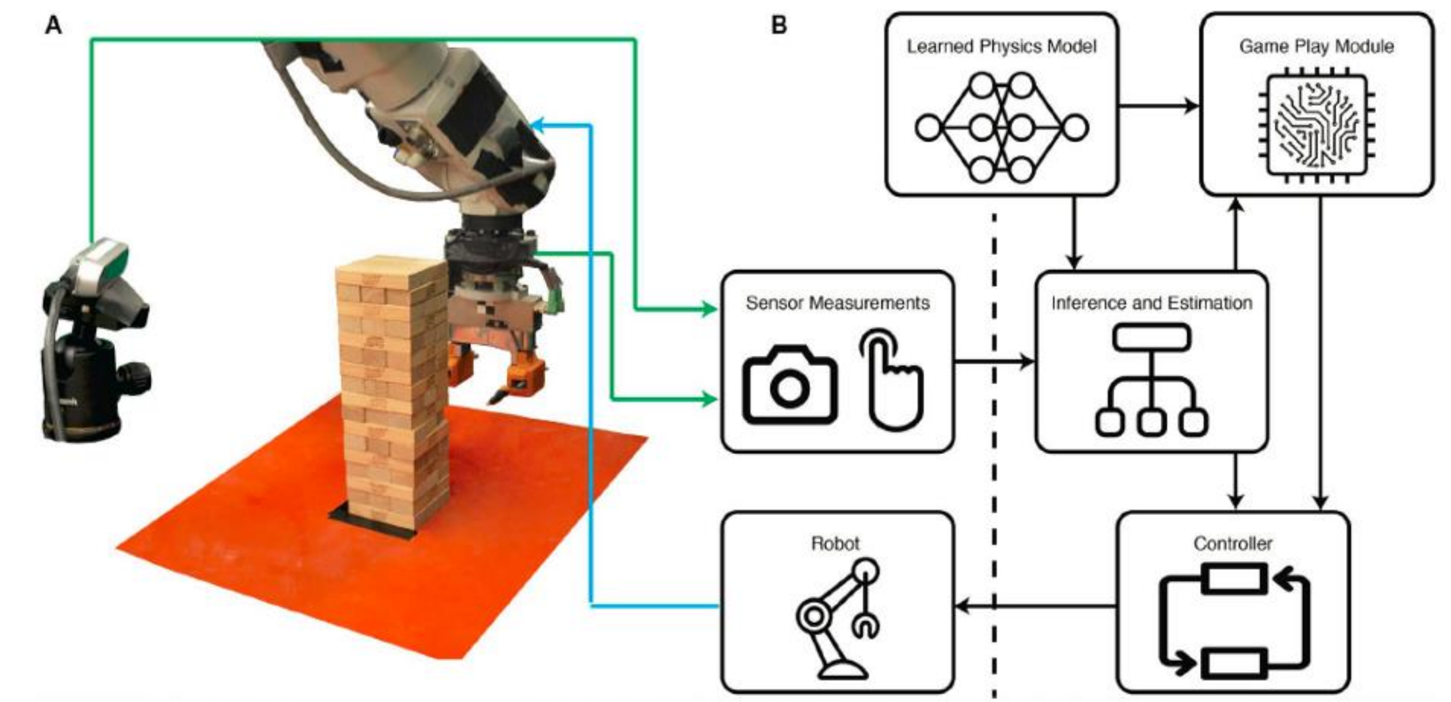


Fig. 1. Robot setup. (A) Physical setup consisting of the robot, Jenga tower, Intel RealSense D415 camera, and ATI Gamma force/torque sensor (mounted at the wrist). (B) Machine intelligence architecture with the learned physics model.

Fig. 4:Experimental setup.

HOW TO PLAY JENGA AS A ROBOT

- Concept learning
- Probabilisitic inference
- Greedy MPC
- Visual system
  - Convolutional network
  - Hidden Markov Model (HMM)

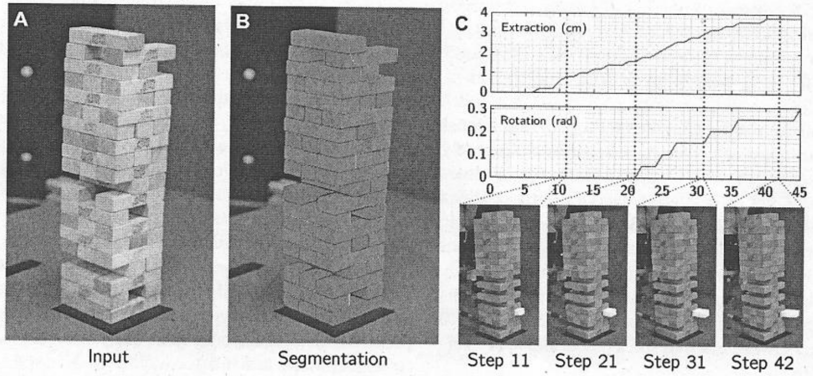


Fig. 7. The vision system. (**A**) We use a 280-by-280 patch with the tower at the center. For clarity, in this figure, we crop the irrelevant regions on both sides. (**B**) Segmented blocks. (**C**) For a single push, we identify the block being pushed and, based on its associated segment, infer the underlying pose and position of the block using an HMM.

Fig. 5: Vision system.

HOW TO PLAY JENGA AS A ROBOT

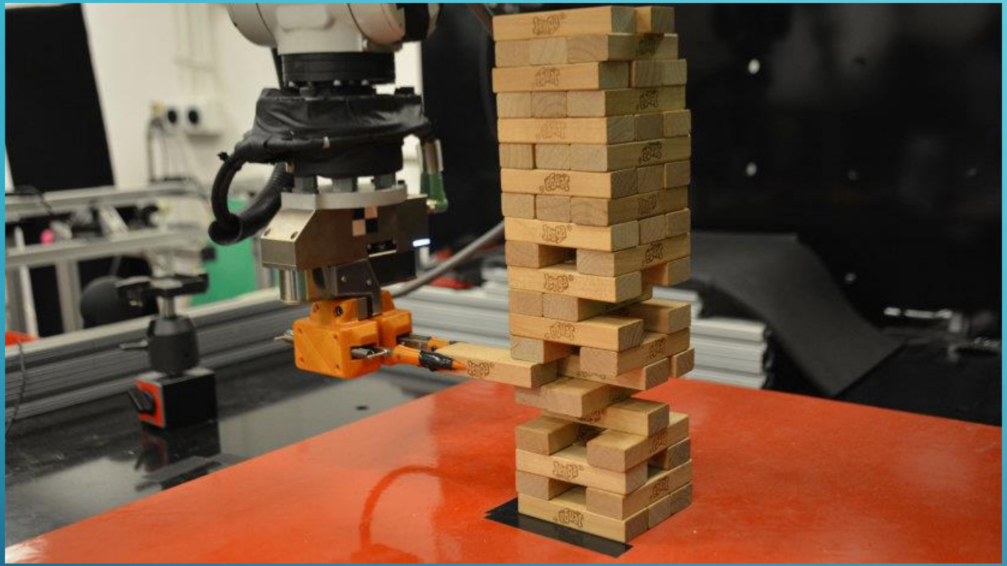


Fig. 6: The robot extracts a block.

- Abstractions via top-down approach
- State-transition via bottom-up approach
- Bayesian neural network (BNN)

HOW TO PLAY JENGA AS A ROBOT

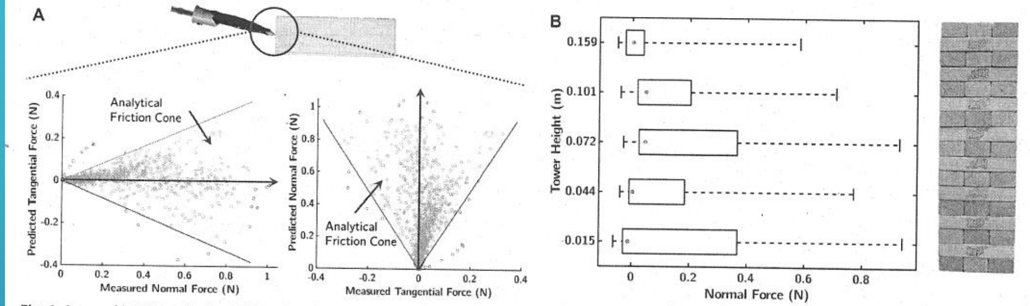


Fig. 4. Learned intuitive physics. (A) Overlay of the analytical friction cone and predicted forces given the current measurements. The friction coefficient between the finger material (PLA) and wood is between 0.35 and 0.5; here, we use 0.42 as an approximation. (B) Normal force applied to the tower as a function of the height of the tower. Each box plot depicts the minimum, maximum, median, and standard deviation of the force measures.

Fig. 7: Analyzing the force.

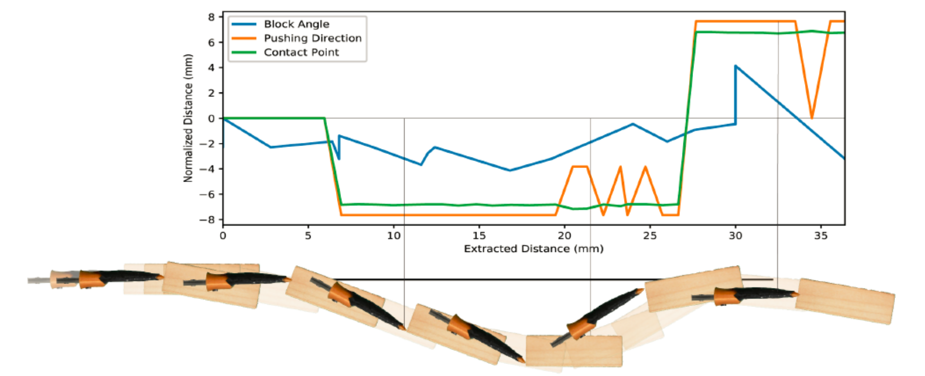


Fig. 6. Controlled block pushing. The robot selects the point on the block and the appropriate angle to push with such that it realigns the block with the goal configuration. Here, the block is beginning to rotate counterclockwise and is starting to move out of the tower. The robot selects a point close to the edge of the block and pushes it back in toward the tower center. We convert angles to normalized distances by scaling with the radius of gyration of the block (0.023 m). We have exaggerated the block translation to illustrate the fine-grained details of motion.

Fig. 8: Block trajectory

- Find the right force
- Block probability
- Abstractions in a control or decision-making framework
    - Action penal

LET'S PLAY JENGA

Table 1. Summary statistics for exploration and learned physics.
A comparison of the performances of the robot using the exploration strategy and the learned model.

Block position	Action	Exploration		Learned	
		Attempts	Successes	Attempts	Successes
All	Push	403	172 (42.7%)	203	96 (45.8%)
	Extract	172	97 (56.4%)	93	82 (88.2%)
	Place	97	85 (87.6%)	82	72 (87.8%)
Side	Push	288	122 (42.4%)	133	69 (51.9%)
	Extract	122	52 (42.6%)	69	54 (78.3%)
	Place	52	44 (84.6%)	54	49 (90.7%)
Middle	Push	115	50 (43.5%)	70	33 (47.1%)
	Extract	50	45 (90.0%)	33	28 (84.8%)
	Place	45	41 (91.1%)	28	23 (82.1%)

Fig. 9: Strategy statistics.

- Visual and Tactile system
  - Additional criteria came up
  - Provided noisy approximations
- The performance results
- Sources for failure

LET'S PLAY JENGA

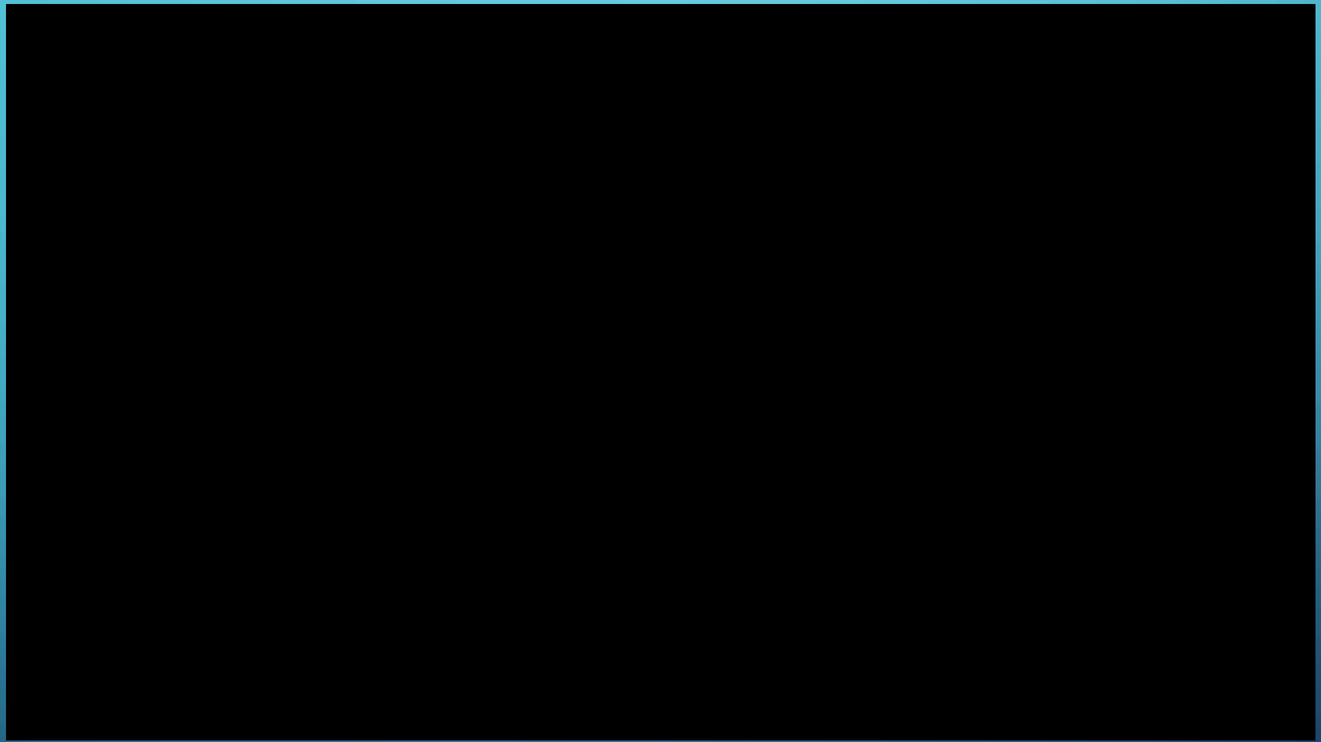
CONCLUSION

- Generalization can be a fruitful endeavor
- Using more than one sense can…
  - …improve the learning rate
  - …tackle new and old issues
  - …be challenging in combining them
- Useful for industrial purposes
  - Electronic assemblage, logistics, disaster response, etc.
- Great example on how games can be used as a testing ground

REMAINING QUESTIONS

- What about the extract/place primitive?
- What if the blocks weren‘t uniform?
- What if a movable camera was used?

THANKS FOR YOUR ATTENTION!



If you have some time to spare watch someone play Jenga in Just Cause 3 ;)
All he has is the visual input and his own knowledge about physics.

REFERENCES AND TABLE OF FIGURES

- Fazeli, N., Oller, M., Wu, J., Wu, Z., Tenenbaum, J. B., Rodriguez, A. (2019): „See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion“, in Science Robotics, Vol. 4, Issue 26.
- https://openai.com/blog/openai-baselines-ppo/
- http://news.mit.edu/2019/robot-jenga-0130
- Fig. 1: https://newatlas.com/mit-jenga-playing-robot/58276/
- Fig. 2: https://robotics.sciencemag.org/content/4/26
- Fig. 3, 5, 7, 9: Fazeli, N., Oller, M., Wu, J., Wu, Z., Tenenbaum, J. B., Rodriguez, A. (2019): „See, feel, act: Hierarchical learning for complex manipulation skills with multisensory fusion“, in Science Robotics, Vol. 4, Issue 26.
- Fig. 4, 8: https://medium.com/syncedreview/ai-gaming-gets-physical-mit-robot-plays-jenga-e911fbaebdff
- Fig. 6: https://entertainment.howstuffworks.com/leisure/traditional-games/how-to-play-jenga-like-youve-never-played-it-before.htm
- Video about the Jenga playing robot: https://www.youtube.com/watch?v=o1j_amoldMs
- Video about Just Cause 3: Jenga!: https://www.youtube.com/watch?v=V6KvekkS7so