ConSOR: A Context-Aware Semantic Object Rearrangement Framework for Partially Arranged Scenes

Kartik Ramachandruni, Max Zuo, Sonia Chernova

OCTOBER 1 - 5, 2023
IEEE/RSJ International Conference on Intelligent Robots and Systems

Georgia Tech

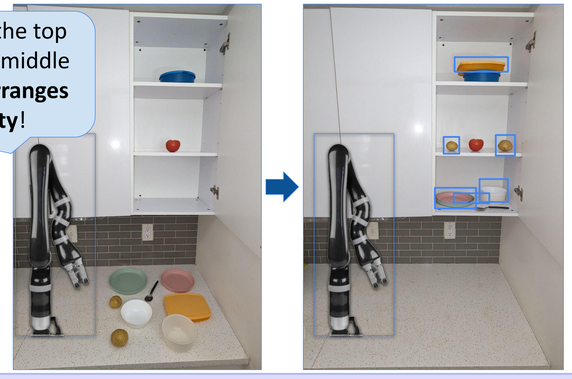
Rearrangement in Partially Arranged Environments

- Household environments are typically pre-arranged with objects.
- This state of *partial arrangement* provides valuable clues regarding the user's organizational preferences (figure on the right).

Motivation

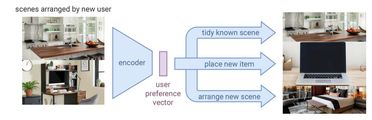
I see a **food container** on the top shelf and a **tomato** on the middle shelf…. I think this user **arranges objects by functionality**!

Our Key Idea: Contextual cues from partially arranged scenes can replace user instruction!



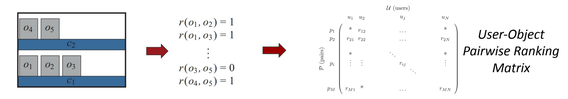
Prior work

- Requires the user to provide explicit instructions or a task demonstration.
- This **burdens the user to communicate their preferences.**



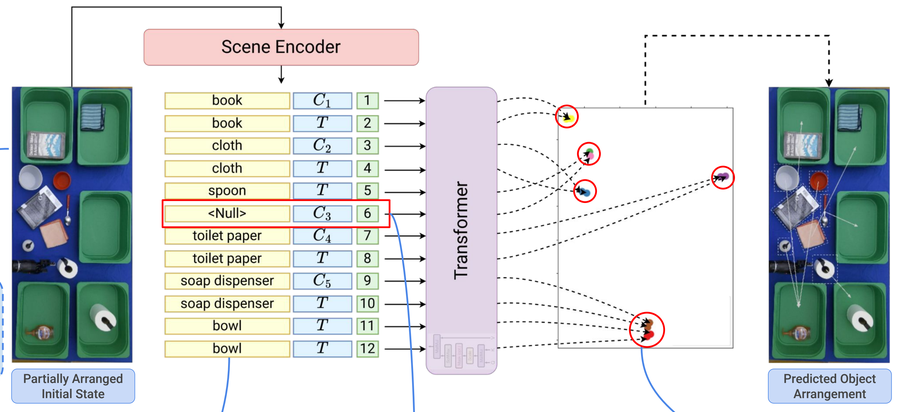
Kapelyukh 2022

Shridhar 2022



Abdo 2016

ConSOR: Context-Aware Semantic Object Rearrangement



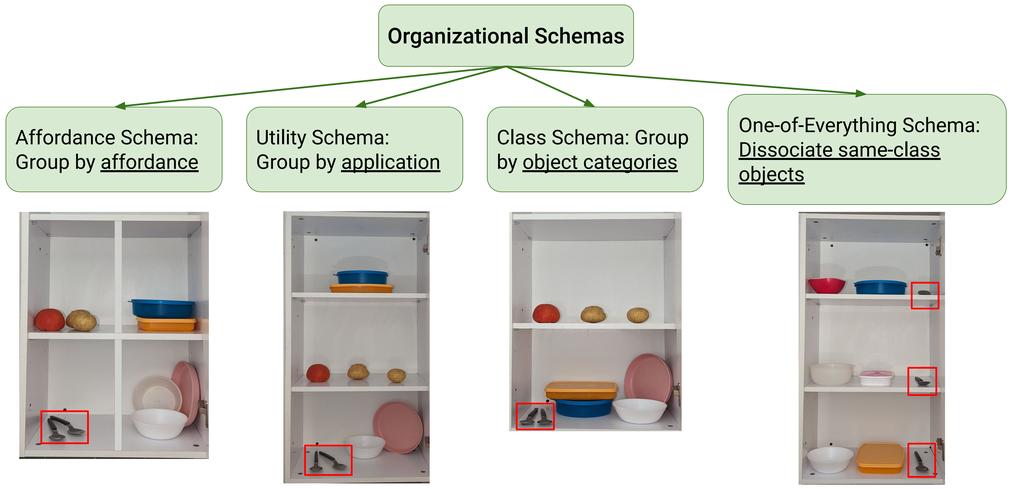
ConSOR relies on **contextual cues from the initial state** such as the placement of pre-arranged objects and environment configuration.

Encoding commonsense knowledge about object categories using ConceptNet enables **generalization to novel objects**.

Our method can also rearrange objects in **environments with empty receptacles**.

Object placements are determined by clustering the latent space.

Dataset of Organizational Schemas for Object Rearrangement



- Our dataset consists of **8800 arranged scenes** using **38 household objects** from AI2Thor.
- The data is divided into scenes using **28 seen and 10 unseen object categories**.

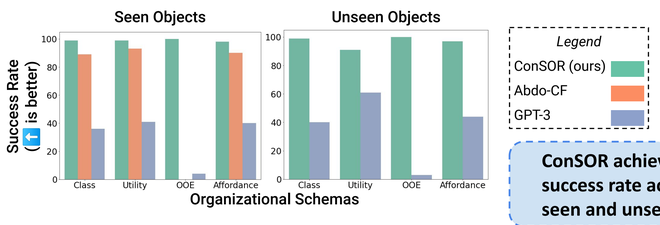
Evaluation Metrics

- **Success Rate**: Measures how often the robot places objects according to the correct organizational style.
- **Non-zero Scene Edit Distance (NSED)**: Measures how many objects the robot misplaces on average when it *incorrectly* predicts the object configuration.

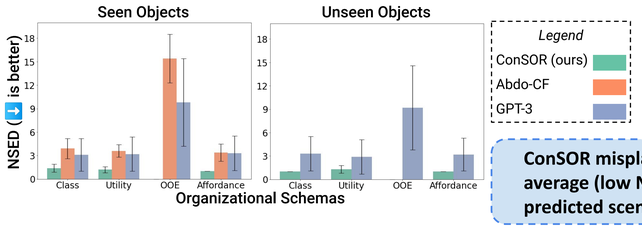
Examples of predicted object arrangements by ConSOR.

Evaluation Results

Evaluating generalization to **Unseen Object Arrangements** and **Novel Object Categories**



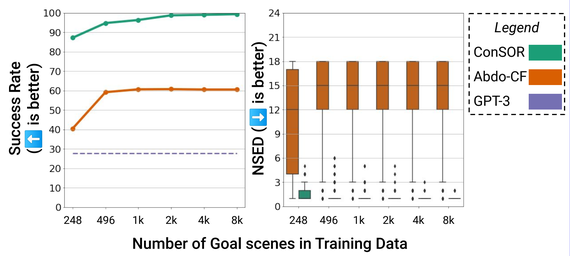
ConSOR achieves the highest success rate across scenes with seen and unseen objects.



ConSOR misplaces less objects on average (low NSED) among incorrectly predicted scenes than baselines.

Baselines

- **Abdo-CF**: Matrix factorization approach that models pairwise object preferences for each schema. Abdo-CF requires schema label as an input.
- **GPT-3**: GPT-3 model prompted with unlabeled few-shot demonstrations from each schema.



Our method retains performance even after reducing the size of training data by 32X.

Conclusions and Future Work

We contribute the following in this work:

- A formulation of the object rearrangement problem in partially arranged environments.
- A semantic reasoning framework for object rearrangement that replaces human instruction with contextual cues from pre-arranged environment.
- An object rearrangement dataset consisting of 8800 arranged scenes using 38 household objects, with each scene belonging to one of 4 organizational schemas.

Future Work:

- Personalization to novel users by relaxing the assumption of a closed set of organizational schemas.
- Identifying placement surfaces and pre-arranged object placements from observation data.

RAIL
Robot Autonomy and Interactive Learning

✉: kvr6@gatech.edu
𝕏/🐦: @kartikvrama
Website: kartikvrama.github.io