

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/378342144>

Robo-CSK-Organizer: Commonsense Knowledge to Organize Detected Objects for Multipurpose Robots

Conference Paper · April 2024

CITATIONS

0

READS

174

4 authors, including:



Rafael Hidalgo
Montclair State University

4 PUBLICATIONS 4 CITATIONS

[SEE PROFILE](#)



Jesse Parron
Montclair State University

9 PUBLICATIONS 5 CITATIONS

[SEE PROFILE](#)



Aparna S. Varde
Montclair State University

181 PUBLICATIONS 1,424 CITATIONS

[SEE PROFILE](#)

Robo-CSK-Organizer: Commonsense Knowledge to Organize Detected Objects for Multipurpose Robots

Rafael Hidalgo¹, Jesse Parron¹, Aparna S. Varde¹, and Weitian Wang¹

School of Computing, Montclair State University (MSU), NJ, USA
`{hidalgor2,parronj1,vardea,wangw}@montclair.edu`

Abstract. In the rapidly evolving field of robotics, integration of commonsense knowledge (CSK) in AI systems is becoming highly crucial to enhance the decision-making capabilities of robots, especially in next-generation multipurpose environments. This paper presents Robo-CSK-Organizer, a pioneering system that employs CSK, via a classical knowledge base, to facilitate sophisticated task-based object organization helpful in multipurpose robots. Unlike systems relying solely on deep learning tools such as ChatGPT, our Robo-CSK-Organizer system stands out in various crucial aspects. This includes: (1) its ability to resolve ambiguities and maintain consistency in object placement; (2) its adaptability to diverse task-based classifications; and moreover, (3) its contributions to explainable AI (XAI), consequently helping to foster trust and human-robot collaboration. This system's efficacy is underlined by DETIC (DEtector with Image Classes), an advanced extension of Detectron2 for object identification; BLIP (Bootstrapping Language-Image Pre-training) for context discernment; and most vitally by the adaptation of ConceptNet, a well-grounded commonsense knowledge base for reasoning based on semantic as well as pragmatic knowledge. While we deploy ConceptNet to extract CSK, the process in Robo-CSK-Organizer is generic enough to be replicated with other state-of-the-art knowledge bases. Controlled experiments and real-world applications, synopsized in this paper, make Robo-CSK-Organizer demonstrate superior performance in placing objects in contextually relevant locations, highlighting its clear capacity for commonsense-guided decision-making closer to the thresholds of human cognition. Hence, Robo-CSK-Organizer makes valuable contributions to Robotics and AI.

AI-Robotics Bridge, Commonsense Reasoning, Explainable Models, Multipurpose Robots, Next-Generation AI Systems, Task Classification

1 Introduction to Robo-CSK-Organizer

The expanding role of multipurpose robots necessitates the development of transparent, intelligent systems capable of handling complex, context-driven operations. Traditional robotic arms, while proficient in assembly tasks, often struggle with nuanced decision-making in dynamic environments. Our work addresses

this challenge by introducing Robo-CSK-Organizer (see Fig. 1), a system that demonstrates Explainable AI (XAI) via Commonsense Knowledge (CSK) to enhance the decision-making capabilities of robots. Imbibing CSK through the ConceptNet knowledge base, this system aims to address the opaque-box issue in AI (typically found in pure deep learning systems), characterized by a lack of decision-making clarity — a growing concern in the realm of robotics where precision and trust are paramount (32).

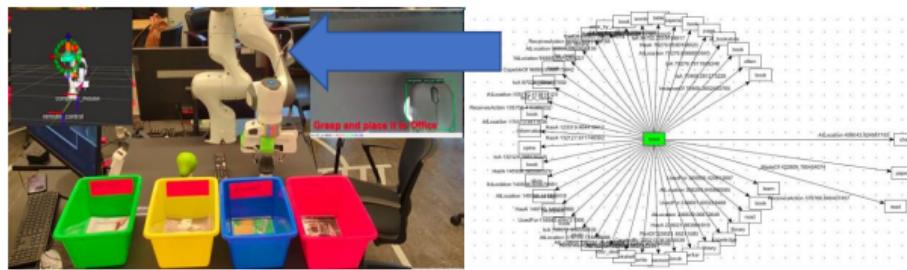


Fig. 1: Graphical abstract of Robo-CSK-Organizer

The Robo-CSK-Organizer system is proposed to address the limitations of current AI systems, which often fall short in tasks requiring deep contextual awareness. Traditional AI models may often struggle with differentiating between similar objects in varied contexts, such as distinguishing a child's toy from a pet's toy, or understanding the appropriate placement of objects in household settings. Robo-CSK-Organizer bridges this gap by adequately deploying commonsense knowledge (CSK) to enhance the organization of detected objects for task classification in multiple avenues.

Note that CSK differs from encyclopedic knowledge (20), (16) (where AI systems far surpass humans). “Common sense” is naturally found in humans who acquire it inherently at birth, enhance it with further growth, and use it for intuitive reasoning. Machines on the other hand are not endowed with CSK by default and hence can often find it challenging to conduct reasoning intuitively unless pre-programmed with rigorous training (9). For instance, it is very easy and in fact rather obvious for a human to know that the “door” of a *refrigerator* should not remain open (except while placing things in it or taking them out) (4). Conversely, the “door” of an *office* can certainly be open and is often ajar - a related fact also quite obvious to humans. Such knowledge is very subtle and is thus considered really simplistic or too common! Yet it can often be crucial in decision-making. Therefore, the role of CSK can be vital in modern-day AI systems as noticed significantly in the literature (33), (1).

Robo-CSK-Organizer precisely addresses this concept of CSK in AI. More specifically, it employs a robust semantic network from a classical knowledge base, namely ConceptNet (26), hence enabling robots to make more well-informed decisions based on contextual cues. This is in line with the logic of harnessing

commonsense knowledge to augment machine intelligence (29). It thrives on the adequate extraction and compilation of CSK from a knowledge base, which is a non-trivial task (23), and can be crucial in AI applications.

2 The Robo-CSK-Organizer Approach

The system diagram of Robo-CSK-Detector, illustrative of its functioning approach, appears in Fig. 2 Its salient features are as described below.

1. **Resolving ambiguity in object categorization:** Classification of objects, such as deciding whether a pear belongs in the kitchen or the garden, highlights the ambiguity inherent in object categorization today. This issue extends to the difficulty of providing comprehensive labeled training examples for every possible object arrangement and is effectively achieved by Robo-CSK-Organizer by harnessing CSK in a task-relevant manner.
2. **Maintaining consistency in object placement:** Ensuring consistency in placing objects is crucial, particularly to build trust among users (e.g. helpful in human-robot collaboration) as well as among other robots in a multi-robot environment. Robots must reason with basic commonsense as well as domain knowledge and manage contextual variations to in order to ensure reliable object placement. This is well-achieved by Robo-CSK-Organizer with clear reasoning paths.
3. **Depicting task relevance and adaptability:** Robots demonstrating adaptability in prioritizing tasks based on context, e.g. choosing between gardening and culinary activities, is vital. This is especially with respect to probabilistic uncertainties in sensing and navigation. Robo-CSK-Organizer handles this very well due to its systematic approach guided by a well-grounded knowledge base.
4. **Fostering explainability in AI systems:** A critical aspect of XAI (Explainable AI) is ensuring that AI systems are not just intelligent but also comprehensible and interpretable. Robo-CSK-Organizer excels in explainability, surpassing other systems (e.g. ChatGPT-based organizers) by adequately harnessing CSK due to which decision-making processes are more transparent and understandable.

In Fig. 3, Robo-CSK-Organizer's utilization of ConceptNet for object categorization in a kitchen setting is visually depicted. This figure illustrates the decision-making pathway for categorizing a pear, among other items. Specifically, it shows 3 potential paths from "kitchen" to "pear". The system selects the path with the highest "AtLocation" edge weight (in this case, 7.21), indicating the common location of food in a kitchen. This path is further delineated by linking "apple" to "food" (via a "RelatedTo" edge) and subsequently connecting "apple" to "pear". This logical sequence leads Robo-CSK-Detector to place the pear in the kitchen, exemplifying the system's reasoning process.

Designed with modules for object detection, context recognition, and semantic analysis, Robo-CSK-Organizer not only classifies objects but also interprets

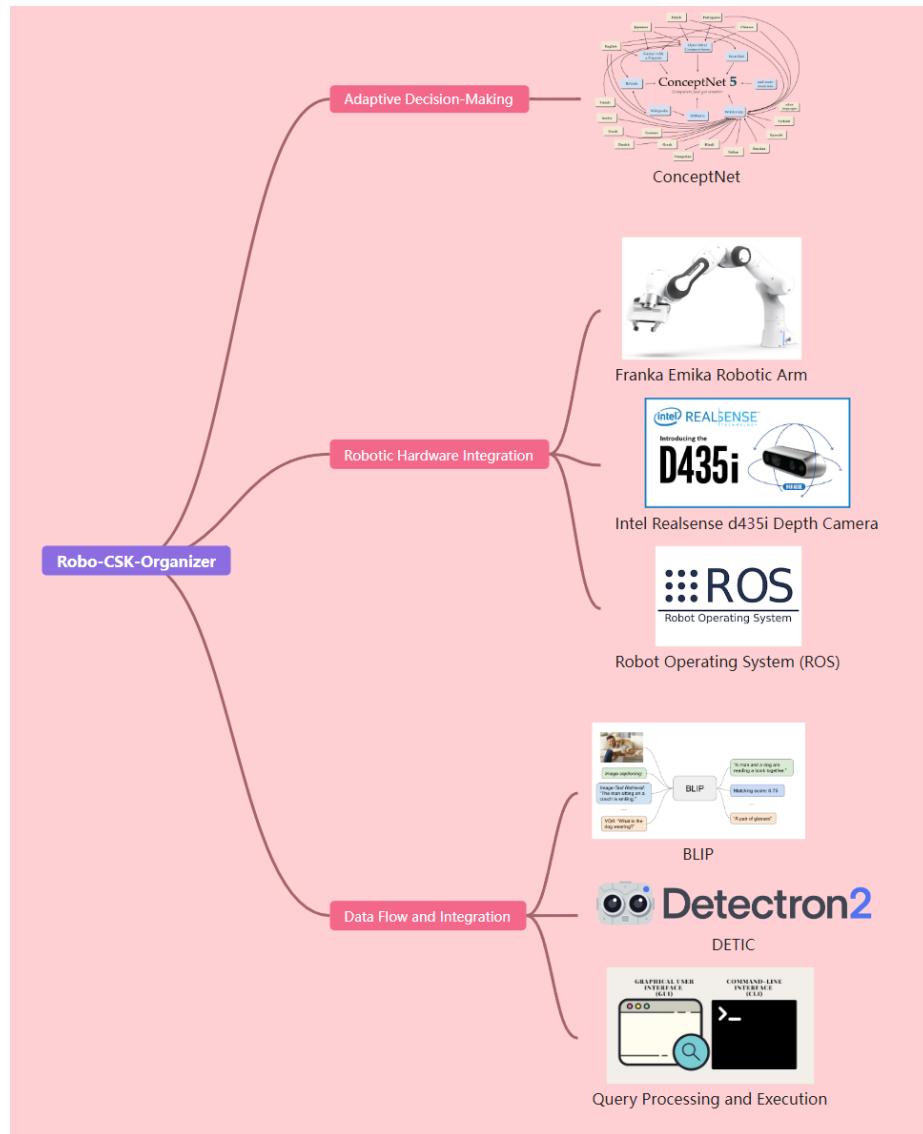


Fig. 2: System diagram of Robo-CSK-Organizer

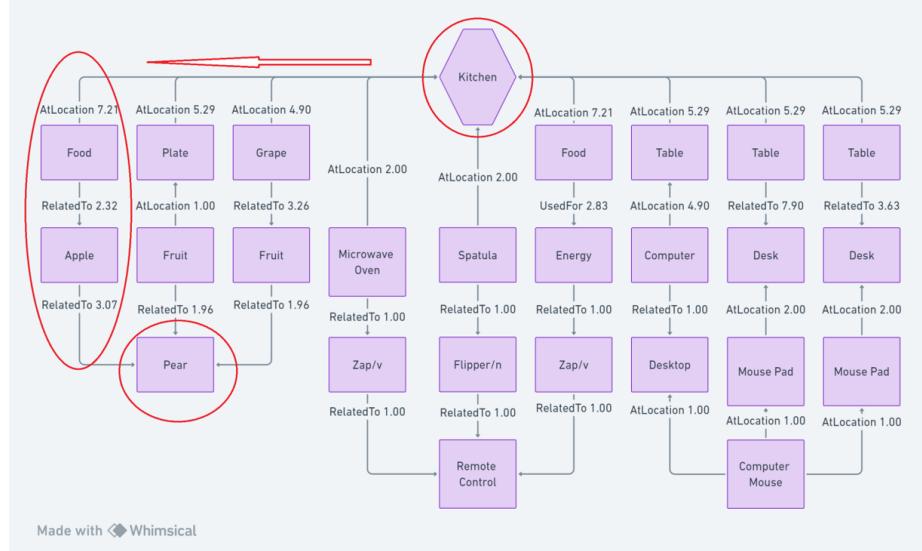


Fig. 3: ConceptNet-based decision pathway in Robo-CSK-Organizer. Illustration of the object categorization process for a pear, highlighting the selection of the optimal path based on 'AtLocation' edge weight and relational connections between 'food', 'apple', and 'pear'

their appropriate placement within various contexts. This approach significantly enhances the transparency and interpretability of AI decisions, addressing the critical challenge of explainability in AI systems. It can thus be considered analogous to systems that delve into aspects such as spatial commonsense (10), (11), (22) for object recognition and training of autonomous systems. For instance, if there is an error, its traceability is highly facilitated in Robo-CSK-organizer (versus deep learning based systems, e.g., those using ChatGPT for training). Hence, Robo-CSK-Organizer can help robots learn from their mistakes and correct themselves, thus getting better in their performance. Moreover, the XAI contribution of Robo-CSK-Organizer is helpful when humans and robots work together, i.e. for human-robot collaboration, as humans are able to understand the actions of the robots much better, along with the reasons behind the robots' decisions. The same logic applies to numerous robots working together. This enhances trust in the realm of robotics. All these facets are vital, especially with the growing prevalence of multipurpose robots, heading towards next-generation advancements.

The main functioning of Robo-CSK-Organizer, focusing on its reasoning, is outlined in Algorithm 1 here.

This algorithm that highlights the main functioning of Robo-CSK-Organizer operates on 2 primary inputs: a video feed \mathcal{V} and the ConceptNet knowledge base \mathcal{C} . Its goal is to sort detected objects into their appropriate contexts. Initially, the robot's vision system \mathcal{R} , implemented using Detectron2, is initialized.

Algorithm 1 Robo-CSK-Organizer Reasoning

Require: \mathcal{V} (Video feed), \mathcal{C} (ConceptNet knowledge base)
Ensure: $\mathcal{O}_{\text{sorted}}$ (Objects sorted into appropriate contexts)

- 1: Initialize robot vision system $\mathcal{R} \leftarrow \text{Detectron2}$
- 2: Scan context bins using $\mathcal{B} \leftarrow \text{BLIP}$, store context in \mathcal{S}_{csv}
- 3: **for** $f \in \mathcal{V}$ **do**
- 4: $\mathcal{D} \leftarrow \text{Detect and label objects in } f \text{ using } \mathcal{R}$
- 5: **for** $o \in \mathcal{D}$ **do**
- 6: $\mathcal{K}_o \leftarrow \text{Query } \mathcal{C} \text{ for context of } o$
- 7: **if** \mathcal{K}_o matches context in \mathcal{B} **then**
- 8: Place o in matched context bin
- 9: **else**
- 10: Continue to next object
- 11: **end if**
- 12: **end for**
- 13: **end for**
- 14: Optional: Display annotated frame from f

Thereafter, the context bins are scanned and recognized using BLIP (\mathcal{B}), with the contexts stored in a CSV (comma-separated variable) format (\mathcal{S}_{csv}). For each frame f in the video feed \mathcal{V} , objects are detected and labeled as \mathcal{D} . Each detected object o is then checked against \mathcal{C} to determine its context \mathcal{K}_o . If this context matches that of a bin in \mathcal{B} , the object is placed in the corresponding bin. The process continues for each object in the frame, and the annotated frame can be displayed as needed.

This algorithm is implemented into the Robo-CSK-Organizer system using Python and is integrated with a robotic arm in our laboratory, namely, the CRoSS Lab (Collaborative Robotics and Smart Systems Lab at our university). Robo-CSK-Organizer is then executed using various real-world objects. Details of its execution are mentioned next in the respective parts of its system demonstration.

3 System Demo and Evaluation

In order to demonstrate the efficacy of Robo-CSK-Organizer, it is compared with a baseline task organizer that uses the well-known ChatGPT for guidance. We thus present the following.

Object Detection: Both the systems, our Robo-CSK-Organizer and the ChatGPT baseline, use DETIC ("DEtector with Image) ensure a broad evaluation spectrum. Specific context groups, particularly domestic locations (e.g. kitchen, garden, pantry, dining room) are chosen for evaluation. These contexts are relevant to the selected object categories and provided with a controlled environment for testing. An advanced extension of Detectron2 for object detection, which contains over 21,000 classes (31) (35) is used here to provide choices of classes for object organization. Each object is queried against each system

(Robo-CSK-Organizer / ChatGPT) 10 times, asking it to organize the object into one of the provided contexts. Responses from both systems are recorded for each iteration, and the most frequent context is identified as the predominant choice for object placement,

Context Recognition: BLIP (Bootstrapping Language-Image Pre-training) is employed to identify contexts such as the kitchen or office, and thus generate room captions for enhanced clarity (19). Note that the usage of such software can be helpful in a variety of applications, e.g. image personalization via text by harnessing diffusion models (14). The hardware foundation for both systems (i.e. Robo-CSK-Organizer and the ChatGPT baseline) includes a Franka Emika robotic arm (12) with an Intel Realsense D435i camera (15), integrated with ROS. Robo-CSK-Organizer works with this hardware, and incorporates CSK-based reasoning (See Algorithm 1). This is the key to addressing the opaque-box issue in AI, aiming for clearer and more transparent object sorting.

Note that the pivotal distinction between the 2 systems lies in the functioning approach for sorting objects into relevant contexts. While ChatGPT relies solely on prior training with deep learning, Robo-CSK-Organizer applies commonsense knowledge due to which it can be more adept in successfully handling first-time scenarios as well. More details appear next.

Robo-CSK-organizer: It harnesses a classical knowledge base called ConceptNet (26) for semantic insights and commonsense reasoning. It infers object locations using metrics known as *edge weight* and *degree of separation*, prioritizing paths based on these factors. While ConceptNet is chosen for its user-friendly interface and clear path logic, we claim that other relevant CSK knowledge bases can also be used.

ChatGPT baseline: A ChatGPT-trained organizer is used as a baseline; it relies on generative pre-trained transformer models for its decision-making, processing text-based inputs to infer object locations and categorizations. This approach, though adept in language processing, does not integrate a structured commonsense knowledge base. Consequently, the ChatGPT-based organizer's decisions are more influenced by pre-trained patterns in textual data rather than explicit semantic relationships and intuitive logical reasoning. This can affect consistency and transparency in decision-making in complex or ambiguous scenarios, notably (but not limited to) those encountered for the first time.

In our comprehensive evaluation, we conduct experiments to assess the performance of Robo-CSK-Organizer and the Chat-GPT baseline across various contexts. The key aspects of these experiments are focused on ambiguity resolution, consistency, task-relevance adaptability, and explainability. A summary of our exhaustive experimentation is presented below.

3.1 Ambiguity Resolution

Both systems are tested on their ability to resolve ambiguous contexts using a variety of objects. Robo-CSK-Organizer as well as ChatGPT baseline performances are evaluated against a ground truth established by semantic similarity scores from state-of-the-art paradigms such as FastText, Word2Vec, and GloVe

models that can be widely accepted as gold standards. The results (See Fig. 4), show that Robo-CSK-Organizer has notable accuracy.

3.2 Ensuring Consistency

The consistency experiments aim to evaluate the stability and repeatability of Robo-CSK-Organizer versus the ChatGPT baseline when faced with identical queries across multiple iterations. This measure of consistency is vital for reliable knowledge organization systems, as it reflects the systems' ability to consistently choose the same context for an object through numerous trials. In the methodology, objects from various categories (e.g. personal items, clothing, office supplies, and toys) are selected. Robo-CSK-Organizer achieves 100% consistency rate across all object-location pairs; this can be attributed to the static nature of the ConceptNet knowledge graphs that it utilizes in its decision-making. In contrast, the ChatGPT Organizer displays less consistency, particularly for objects such as adhesive tape, belt, sock, remote control, toothpaste, and aerosol can; possibly indicating that pre-training alone may not always yield consistent results in systematic object organization.

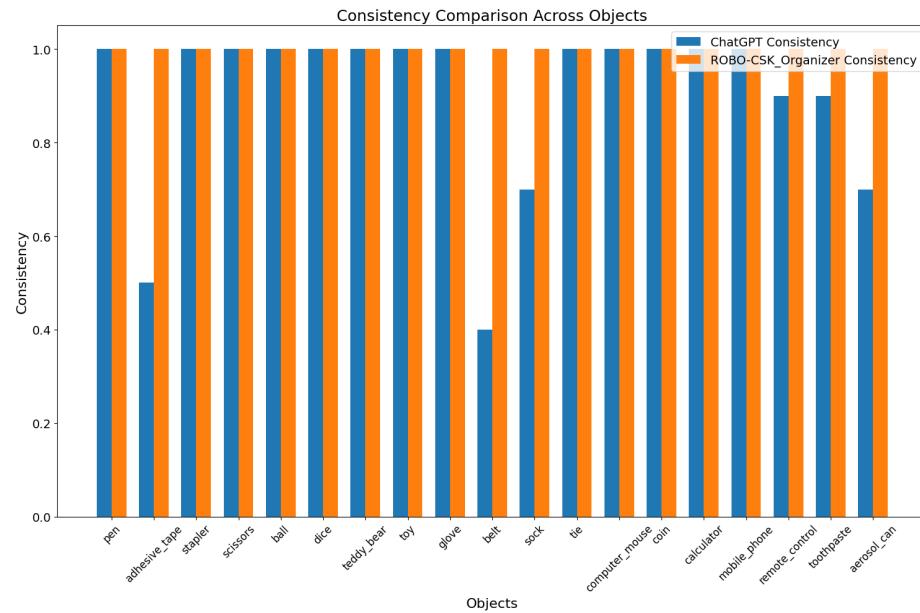


Fig. 4: Robo-CSK-Organizer has 100% consistency in all object-location pairs. ChatGPT baseline is not as consistent for all objects, specifically for adhesive tape, belt, sock, remote control, toothpaste, and aerosol can

3.3 Task-Relevance Adaptability

These experiments evaluate the systems' ability to adapt their responses to different context-specific directives. The objective is to assess whether Robo-CSK-Organizer and the ChatGPT baseline can re-calibrate their responses when directed to focus on alternative contexts differing from their initial preferences. The experiments commence with an initial straightforward assessment, querying "apple" against 4 contexts (kitchen, living room, bedroom, bathroom) without any focused directives. This determines the systems' natural inclination or preference for context association. In the adaptability testing phase, the systems are prompted to focus on the remaining 3 contexts, one at a time, to observe if they can adapt their responses when a specific context is emphasized. This is analogous to humans adapting to different contexts in real life, e.g. if you specifically tell a human not to place apples in a kitchen (for any reason, e.g. the kitchen is too small or it is being cleaned for pest control). then the human should intuitively find another good place for the apples rather than placing them in the kitchen again. Accordingly, it is interesting to assess how robotic systems would behave in such situations.

Likewise, data for the initial as well adaptability tests are collected by repeating each context-query 10 times. Responses are compiled into respective data frames for detailed analysis. The focused contexts for the adaptability tests are based on preference, with the most preferred context excluded to emphasize the remaining contexts. The initial phase identifies kitchen as the clear preference for sorting apples for both the ChatGPT-based organizer and Robo-CSK-Organizer. In the adaptability phase, there are observable shifts in Robo-CSK-Organizer's response, as desired, i.e. it is more adaptable when needed. The paths Robo-CSK-Organizer employs for sorting, leading to object placement, are as follows:

- Path: Kitchen (AtLocation) <- food (RelatedTo) > apple
- Path: Bedroom (AtLocation) <- house (AtLocation) <- apple

Figs 5 and 6 here provide a well-summarized visual representation of these findings.

3.4 Explainability

The experiment on explainability assesses the Robo-CSK-Organizer system and the ChatGPT-based system as per their abilities to elucidate their decision-making processes. This aspect is crucial for building user trust and understanding, considering the fact that there are situations where decisions may seem counter-intuitive at times. Robo-CSK-Organizer utilizes Detic for object detection, BLIP for context recognition, and ConceptNet for commonsense knowledge. It can provide logical paths for its decisions, enhancing transparency. For instance, during its incorrect placement of beer into the playroom, Robo-CSK-Organizer provides a clear logical path: *playroom (UsedFor) fun (RelatedTo)*

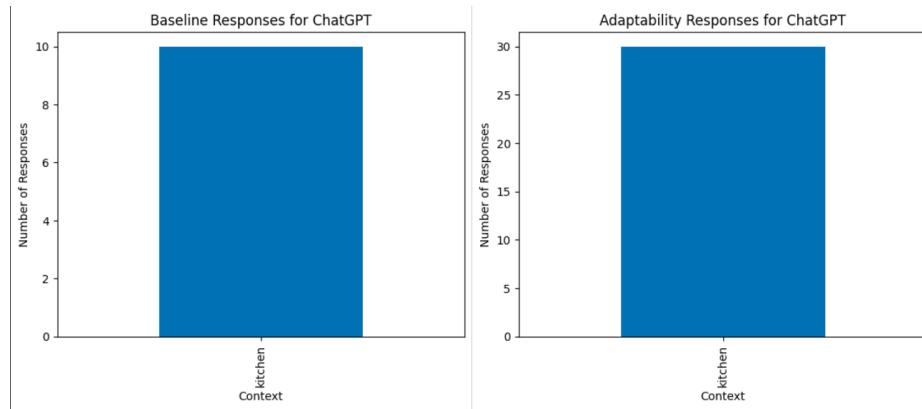


Fig. 5: In the adaptability phase, ChatGPT's implementation does not cause any observable shifts, despite requesting it to change its context.

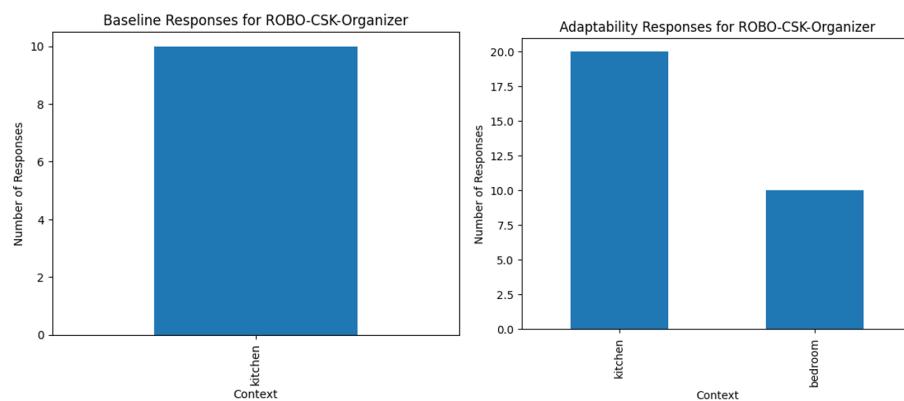


Fig. 6: There are observable shifts on the Robo-CSK-Organizer in one of the locations, after requesting context-switching.

party (RelatedTo) beer. This path demonstrates the connection of concepts leading to the system's conclusion. In contrast, the ChatGPT-based organizer, relying basically on deep learning models, functions as an opaque-box. It is unable to ascertain the explicit reasoning behind its decisions, e.g. placing "scissors" into a "playroom" (which can be a potentially hazardous decision). Hence, the ChatGPT baseline is lacking in a clear explainable framework. This can pose problems in error-correction, thus adversely impacting performance.

This distinction highlights that while both systems may err, analogous to the adage "*to err is human*", Robo-CSK-Organizer's explainability allows for better understanding and correction of these errors. Explainability is essential, especially in robotic systems where precision and safety are critical, contributing to user trust and understanding of AI decisions. Note that such explainability can in turn help in explicit communication with various AI systems, including intelligent agents in mobile apps, e.g. it can help to enhance existing apps with virtual voice agents (17) by adding more image-based functions where adequate object recognition is crucial. Hence, it can indirectly help a different type of robot, including a chatbot or a virtual voice assistant. All these systems would benefit from easier comprehension and enhanced interpretability. Hence, explainable AI plays a vital role here.

Focusing on such aspects, Figs 7 and 8 illustrate how Robo-CSK-Organizer and the ChatGPT-based organizer derive their decisions in the placement of "scissors". The comparative analysis emphasizes the Robo-CSK-Organizer's strengths in consistency, adaptability, and explainability. The findings of our experimentation thus underscore the importance of integrating structured knowledge bases in AI systems. This fact is highlighted here, considering various scenarios for domestic environments.



Fig. 7: Robo-CSK-Organizer's (quite adequate) placement of scissors into the dining room

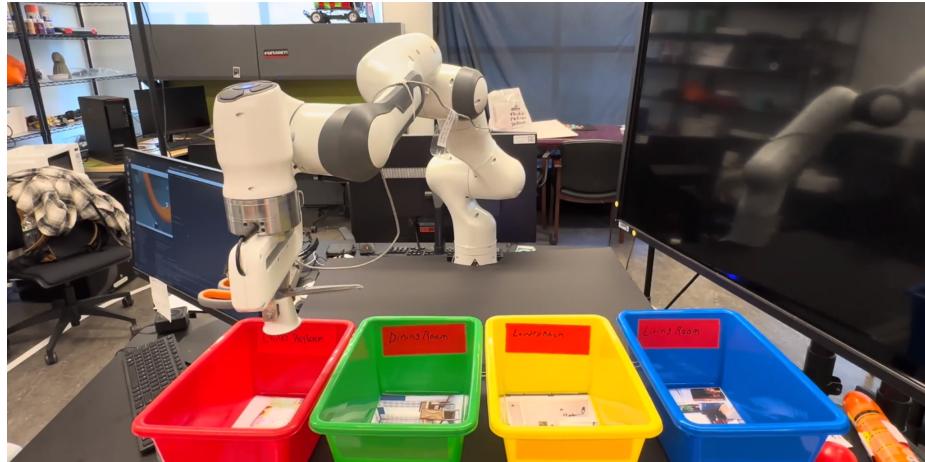


Fig. 8: ChatGPT-Organizer's (rather dangerous) placement of scissors into the child's playroom

Finally, we synopsize the comparative evaluation of Robo-CSK-Organizer and the ChatGPT baseline in a TABLE 1 here.

Table 1: Comparison of Robo-CSK-Organizer and ChatGPT

Parameter	Robo-CSK-Organizer	ChatGPT
Approach	Uses ConceptNet, logical paths	Generative transformer models
Ambiguity Resolution	Noticeable Accuracy	Varies based on trained patterns
Consistency	100% consistent across all classes	Consistency varies depending on class
Task-Relevance Adaptability	Adaptable to directives	Limited adaptability shown
Explainability	High; clear paths	Lower; opaque due to AI
Decision-Making Basis	Semantic, pragmatic CSK	Textual data patterns
Hardware Integration	Robotic systems integration	Robotic systems integration
Use of AI	AI with knowledge bases	Primarily AI-driven

4 Discussion on Main Contributions

This research presents significant contributions to the field of AI and robotics, particularly in the development and application of commonsense knowledge for task classification in multipurpose robots. Robo-CSK-Organizer, as a pioneering system, stands out in several key areas when compared to existing systems (e.g. a baseline organizer using ChatGPT).

4.1 Novel Integration of CSK

Robo-CSK-Organizer's integration of CSK through ConceptNet is a major advancement. Unlike systems that rely primarily on deep learning, Robo-CSK-Organizer utilizes structured knowledge bases in addition to machine learning, to enhance decision-making transparency and accuracy. This integration allows the system not only to recognize and categorize objects but also to self-understand the task contexts, leading to more intuitive and contextually appropriate object operations.

4.2 Superiority in Consistency and Adaptability

Our experiments depict Robo-CSK-Organizer's superiority in consistency and task-relevance adaptability. It achieves a 100% consistency rate, which is crucial for user trust and predictability. Its ability to adapt to different context-specific directives showcases its potential in dynamic and changing environments, which is essential for practical applications such as next-generation multipurpose robots, e.g. those meant to be helpful in domestic settings.

4.3 Advancements in Explainability

Robo-CSK-Organizer makes strides in explainability, a key aspect of XAI (Explainable AI). Its ability to provide logical paths for its decisions enhances user understanding and trust, particularly in situations where decisions might seem counter-intuitive. This feature sets it apart from more opaque-box systems such as those relying solely on deep learning. The more explainable a system is, the easier it is to work with, especially when multiple robots work together and / or humans and robots collaborate with each other.

5 Related Work

The integration of AI and robotics, particularly in domestic environments, has seen a variety of innovative approaches. These methodologies have significantly contributed to the field by enhancing robots' decision-making processes, adaptability, and interaction with their environment.

One approach focuses on using ConceptNet and Google search data for object categorization in domestic robotics, particularly for tidy-up services. This method effectively groups objects into functional categories, thereby aiding robots in more intuitive object handling (25). Another study explores the use of large language models (LLMs) like GPT-3.5 as a repository of CSK for task planning. This demonstrates the potential of language models in enriching the robotic decision-making process (34). Furthermore, paradigms based on the classical neural networks have been adapted to many contexts, ranging from machine translation (7) in text with recurrent neural networks (RNNs) to object recognition in multifaceted scenarios with computer vision models, e.g. VGG-16 (24)

and ResNet-101 (13). The issue of extracting cultural commonsense knowledge and its usefulness in enhancing chatbots has been addressed through a novel approach called CANDLE (21) with interesting real-world impacts.

Advances in visual commonsense reasoning introduce the R2C engine (33) to enhance object recognition, anchoring natural language descriptions in visual data. CSK-Detector (3) is an innovative system for object detection in domestic robotics, leveraging CSK from the Dice knowledge base (2); it reduces the need for extensive image annotation.

The incorporation of CSK from the OMICS database using Description Logic has also been discussed. This integration enables robots to perform more nuanced tasks, showcasing the potential for more context-aware robotics (18). Furthermore, the application of CSK in human-robot collaborative tasks has been highlighted, especially in robot action planning for assembly tasks, emphasizing the enhancement of cooperative interactions (5). Its mathematical modeling insights along with core applications in smart manufacturing have been elaborated (6) as well, emphasizing the crucial role of commonsense reasoning.

Additionally, semantic task planning for service robots in dynamic, open-world environments has been explored. This method leverages natural language understanding and semantic reasoning, addressing the challenges posed by ever-changing environments (8). The combination of non-monotonic logical reasoning and incomplete CSK with inductive learning to guide deep learning in robotics is another innovative approach. This integration offers a unique perspective on the convergence of CSK and advanced learning techniques (27). Much of this work builds upon semantic advances over the years (28), (30) that help in managing knowledge and conducting predictive analysis.

These diverse methodologies underscore the importance of CSK in improving the functionality and intelligence of robotic systems, especially in domestic settings. They have advanced the field by demonstrating effectiveness in task planning, human-robot interaction, and environmental adaptation.

Building on these foundations, the Robo-CSK-Organizer system represents a significant advancement in the practical application of CSK in robotics. Unlike the existing systems, Robo-CSK-Organizer harnesses CSK in real-world settings for object organization in task-based classification, which is particularly beneficial for multipurpose robots. Its ability to resolve ambiguities and maintain consistency in object placement, adaptability to diverse task classifications, and contributions to explainable AI (XAI) set it apart from the current methodologies. This system not only categorizes and understands objects in various contexts but also intelligently organizes them, demonstrating a novel and practical application of CSK in enhancing the efficiency and functionality of multipurpose robotic systems.

6 Conclusions and Roadmap

Robo-CSK-Organizer is a system proposed in this paper that effectively demonstrates XAI via CSK for object organization, mitigating the challenges of non-

transparent AI. Our early experiments open up areas for enhancement, e.g. decision paths from a CSK source such as ConceptNet. For instance, misplacement of high heels in the kitchen can be due to semantic overlap with stiletto as heel / knife. When this is clearly explained, potential improvements can be made. In our work, these can include refining relationships (e.g. RelatedTo) so as to provide better contextual accuracy, and using average weights in the knowledge base, not just the weight of the first edge in order to provide more robustness. Additionally, we are committed to refining the algorithmic logic used to identify the most optimal paths based on CSK.

Impacts of Robo-CSK-Organizer are highlighted here.

1. Puts forth our objective to quantify enhancements that CSK brings to the reliability and transparency of AI
2. Elevates the efficacy of robotic decision-making to bring it closer to human cognition
3. Fosters a broader academic dialogue on commonsense in robots for better interpretation, trust, and explainability
4. Can be useful in next-generation multipurpose robots, and in human-robot collaboration due to higher clarity.
5. Can lead to energy savings due to more efficient learning, thus positively impacting sustainable AI.
6. Well-mounted on an AI-robotics bridge, particularly that of explainable AI and multipurpose robotics.

This paper thus offers our modest contributions to both AI and robotics. We anticipate fruitful, long-lasting impacts.

Acknowledgments

The commonsense knowledge project thrived on a research visit by Dr. Aparna Varde at the Max Planck Institute for Informatics, Saarbrucken, Germany, with further work at Montclair. This work is supported in part by the National Science Foundation under Grants CMMI-2138351 and CNS-2117308. Our experiments are conducted in the CRoSS (Collaborative Robotics and Smart Systems) Lab at Montclair, of which Dr. Weitian Wang is the Director. We thank CESAC (Clean Energy and Sustainability Analytics Center) at Montclair, of which Dr. Aparna Varde is an Associate Director.

Bibliography

- [1] Cambria, E., Liu, Q., Decherchi, S., Xing, F., Kwok, K.: Senticnet 7: A commonsense-based neurosymbolic ai framework for explainable sentiment analysis. In: LREC conf. pp. 3829–3839 (2022)
- [2] Chalier, Y., Razniewski, S., Weikum, G.: Dice: A joint reasoning framework for multi-faceted commonsense knowledge. In: International Workshop on the Semantic Web (2020), <https://api.semanticscholar.org/CorpusID:226263991>
- [3] Chernyavsky, I., Varde, A.S., Razniewski, S.: Csk-detector: Commonsense in object detection. In: IEEE Big Data. pp. 6609–6612 (2022). <https://doi.org/10.1109/BigData55660.2022.10020915>
- [4] Choi, Y.: The curious case of commonsense intelligence. *Daedalus* **151**(2), 139–155 (2022)
- [5] Conti, C.J., Varde, A.S., Wang, W.: Robot action planning by commonsense knowledge in human-robot collaborative tasks. *IEEE IEMTRONICS* pp. 1–7 (2020), <https://api.semanticscholar.org/CorpusID:222298196>
- [6] Conti-C-J., Varde, A.S., Wang, W.: Human-robot collaboration with commonsense reasoning in smart manufacturing contexts. *IEEE Transactions on Automation Science and Engineering* **19**(3), 1784–1797 (2022). <https://doi.org/10.1109/TASE.2022.3159595>
- [7] Corallo, L., Li, G., Reagan, K., Saxena, A., Varde, A.S., Wilde, B.: A framework for german-english machine translation with GRU RNN. In: ACM EDBT workshops. vol. 3135 (2022), https://ceur-ws.org/Vol-3135/darliap_paper4.pdf
- [8] Cui, G., Shuai, W., Chen, X.: Semantic task planning for service robots in open worlds. *Future Internet* **13**(2) (2021)
- [9] Davis, E., Marcus, G.: Commonsense reasoning and commonsense knowledge in artificial intelligence. *Communications of the ACM* **58**(9), 92–103 (2015)
- [10] Garg, A., Tandon, N., Varde, A.S.: I am guessing you can't recognize this: Generating adversarial images for object detection using spatial commonsense. In: AAAI Conf. on Artificial Intelligence. vol. 34, pp. 13789–13790 (2020)
- [11] Garg-A, Tandon, N., Varde, A.S.: CSK-SNIFFER: commonsense knowledge for sniffing object detection errors. In: ACM EDBT workshops. vol. 3135 (2022), https://ceur-ws.org/Vol-3135/bigvis_short2.pdf
- [12] Haddadin, S., Parusel, S., Johannsmeier, L., Golz, S., Gabl, S., Walch, F., Sabaghian, M., Jähne, C., Hausperger, L., Haddadin, S.: The franka emika robot. *IEEE Robotics Automation Magazine* **29**(2), 46–64 (2022)
- [13] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. *CoRR abs/1512.03385* (2015), <http://arxiv.org/abs/1512.03385>
- [14] Hidalgo, R., Salah, N., Chandra Jetty, R., Jetty, A., Varde, A.S.: Personalizing text-to-image diffusion models by fine-tuning classification for ai applications. In: Intelligent Systems Conference. pp. 642–658. Springer (2023)
- [15] Intel Corporation: Intel realsense depth camera d435i: Specifications. Online (2023), available: <https://www.intel.com/content/www/us/en/products/sku/190004/intel-realsense-depth-camera-d435i/specifications.html>
- [16] Joshi, M., Lee, K., Luan, Y., Toutanova, K.: Contextualized representations using textual encyclopedic knowledge. arXiv preprint arXiv:2004.12006 (2020)

- [17] Kalvakurthi, V., Varde, A.S., Jenq, J.: Hey dona! can you help me with student course registration? AAAI Conference, Workshop on AI for Education, arXiv:2303.13548 (2023)
- [18] Kunze, L., Tenorth, M., Beetz, M.: Putting people's common sense into knowledge bases of household robots. In: Dillmann, R., Beyerer, J., Hanebeck, U.D., Schultz, T. (eds.) KI 2010: Advances in Artificial Intelligence. pp. 151–159. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
- [19] Li, J., Li, D., Xiong, C., Hoi, S.: Blip: Bootstrapping language-image pre-training for unified vision-language understanding and generation (2022)
- [20] Nastase, V., Filippova, K., Milne, D.N.: Summarizing with encyclopedic knowledge. In: TAC (2009)
- [21] Nguyen, T., Razniewski, S., Varde, A.S., Weikum, G.: Extracting cultural commonsense knowledge at scale. In: ACM Web Conf. WWW. pp. 1907–1917 (2023)
- [22] Persaud, P., Varde, A.S., Robila, S.: Enhancing autonomous vehicles with commonsense: Smart mobility in smart cities. In: IEEE ICTAI. pp. 1008–1012 (2017)
- [23] Razniewski, S., Tandon, N., Varde, A.S.: Information to wisdom: Commonsense knowledge extraction and compilation. In: Proceedings of the 14th ACM International Conference on Web Search and Data Mining. pp. 1143–1146 (2021)
- [24] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition (2015)
- [25] Skulkittiyut, W., Lee, H., Ngo Lam, T., Tran Minh, Q., Baharudin, M.A., Fujioka, T., Kamioka, E., Mizukawa, M.: Commonsense knowledge extraction for tidy-up robotic service in domestic environments. In: Adv. Robotics (W). pp. 63–69 (2013)
- [26] Speer, R., Chin, J., Havasi, C.: Conceptnet 5.5: An open multilingual graph of general knowledge (2018)
- [27] Sridharan, M., Mota, T.: Towards combining commonsense reasoning and knowledge acquisition to guide deep learning. Autonomous Agents and Multi-Agent Systems **37**(1) (2022)
- [28] Suchanek, F.M., Varde, A.S., Nayak, R., Senellart, P.: The hidden web, xml and the semantic web: Scientific data management perspectives. In: ACM EDBT Conf. pp. 534–537 (2011)
- [29] Tandon, N., Varde, A.S., de Melo, G.: Commonsense knowledge in machine intelligence. ACM SIGMOD Record **46**(4), 49–52 (2018)
- [30] Varde, A.S., Takahashi, M., Rundensteiner, E.A., Ward, M.O., Maniruzzaman, M., Sisson Jr, R.D.: Apriori algorithm and game-of-life for predictive analysis in materials science. International Journal of Knowledge-based and Intelligent Engineering Systems **8**(4), 213–228 (2004)
- [31] Wu, Y., Kirillov, A., Massa, F., Lo, W., Girshick, R.: Detectron2. Online (2019), available: <https://github.com/facebookresearch/detectron2>
- [32] Zednik, C.: Solving the black box problem: A normative framework for explainable artificial intelligence (2019)
- [33] Zellers, R., Bisk, Y., Farhadi, A., Choi, Y.: From recognition to cognition: Visual commonsense reasoning (2019)
- [34] Zhao, Z., Lee, W.S., Hsu, D.: Large language models as commonsense knowledge for large-scale task planning (2023)
- [35] Zhou, X., Girdhar, R., Joulin, A., Krähenbühl, P., Misra, I.: Detecting twenty-thousand classes using image-level supervision (2022)