Building Robots with Crow-Level Tool Use: The Long Road Ahead

Yizhe Huang
School of Intelligence Science and Technology
Peking University
szhyz@pku.edu.cn

Abstract

Abundant evidence highlights the remarkable tool-using abilities of crows, including their capacity to not only utilize tools but also craft them. Yet, we find ourselves situated on an extensive and challenging journey towards equipping machines with the same tool-use proficiency exhibited by crows. This article sets forth a tool-use framework for robots, offering a potential solution to the complex challenge of integrating tools into a robotic super system. However, it is evident that the realization of the functions outlined within this framework necessitates significant advancements in the domains of planning, learning and applying the affordance model, exploration in decision-making and other aspects of work. This article explores the promising path forward in our quest to imbue robots with crow-like tool-use capabilities.

1 The way to crow-level tool use

The use of tools has long been considered a hallmark of human distinctiveness, setting us apart from the rest of the animal kingdom. However, a closer examination of nature reveals that tool use is not solely the domain of humans; other creatures, such as crows, exhibit remarkable tool-using abilities [7, 10]. In [10], crows have been observed using sticks to extract prey from hidden locations, and even fashioning hooks at the end of twigs to create more effective tools. Remarkably, they can transmit these acquired skills to their fellow crows.

The aspiration for an ideal robotic system lies in its ability to utilize tools with flexibility like humans. Yet, even realizing crow-level tool use in robots remains a formidable challenge. [6] have classified the various facets of robotic tool use, as depicted in Fig. 1. Among these classifications, the crow's use of sticks for extracting prey represents an improvisatory tool use, as it involves employing an object not originally designed for such a purpose. The transformation of a twig into a more effective tool represents the epitome of tool creation, categorized as a form of multi-manipulation tool use is a kind of tool manufacturing. [6] has associated the process of tool use with the concept of affordance models. This system covers all causal tool use situations and even some combinations of them.

To achieve improvisatory tool use akin to that of crows, it becomes imperative to comprehend the complete affordance model of the tool. This understanding must encompass not only how to achieve desired outcomes through specific actions but also how to adapt these actions when faced with novel tools sharing similar physical attributes to the learned ones. Furthermore, discerning the precise features of a tool that produce specific effects is of paramount importance. The realm of tool manufacturing introduces additional complexities, demanding more complex manipulation skills. It transcends mere knowledge of the affordance model; it necessitates the ability to modify the affordance of an object through actions, ultimately tailoring it to achieve the desired affordance model.

In the explored realm of deductive tool use and multiple-manipulation tool techniques, current research, as highlighted in [6], is notably lacking in depth, with some areas being completely devoid

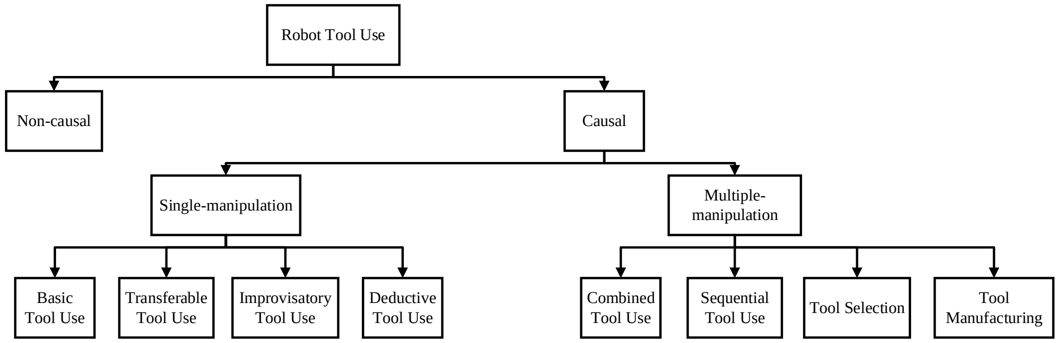

Figure 1: Tool use taxonomy. It is modified from Figure 2 of [6].

of relevant studies. Furthermore, many existing studies in this domain rely on oversimplified versions of these concepts, underscoring the considerable gap that exists between the current state of research and the realization of comprehensive tool-use applications.

2 A goal-directed tool-use system

While [6] offers a classification of diverse tool use scenarios, it falls short of providing a unified framework that encompasses all these scenarios. Moreover, it does not delve into the potential combination of various types of tool use cases. In response to this gap, we propose a framework, which is shown in Fig. 2, that encompasses all the tool-use scenarios illustrated in Fig. 1. A robot based on this framework may become a super system for tool-use.

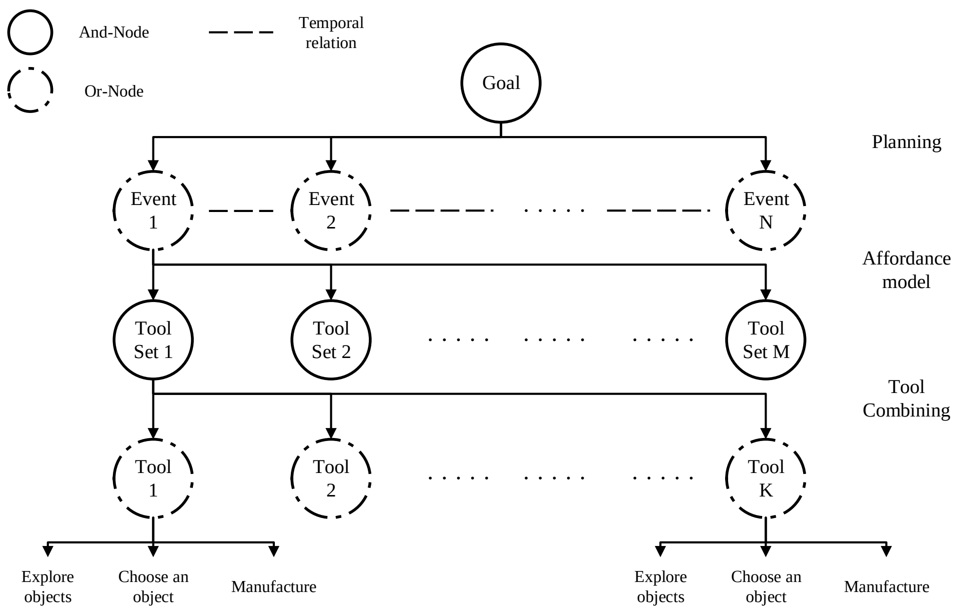

Figure 2: A unified tool-use framework represented as an and-or graph.

First and foremost, it's essential to clarify that we've expanded upon the conventional notion of a "tool" as described in [6]. In our context, a tool is not limited to a specific, predefined object; instead, it encompasses a collection of objects that share a common affordance. A tool can comprise multiple objects, and a single object can serve as a component of different tools.

The system we propose revolves around the achievement of specific goals, aligning with several hypotheses that explain human behavior [2, 1]. This goal may be a partial achievement required to complete a task, or it may be the task itself.

When confronted with a goal, our initial step is to query a planning module to devise a sequence of events required for its attainment. These events may come with distinct temporal constraints. For instance, sequential tool use can be achieve with sequential constraints. When facing with each event, we need to determine the appropriate tool sets to employ, contingent upon our comprehension and learned knowledge of tool affordances. Various single-manipulation tool uses in Fig. 1 are involved here, depending on whether the tool and the event exist in our prior experiences. When facing the same event, we can have multiple tool sets as alternatives. This is where “tool selection" is located. Employing multiple tools simultaneously necessitates skills in “combined tool use". When it comes to acquiring the desired tool (an object with the desired affordance) from the environment, we face the question of how to obtain it. If the choice is made to modify some objects in the surroundings to adapt their affordance and create tools, this process is referred to as “tool manufacturing".

3 Challenges to tackle

There are many hurdles that need to be overcome in order to build such a system introduced in Sec. 2. As it stands, there exists no comprehensive solution for each facet of this ambitious endeavor.

One of the paramount challenges in this pursuit lies in implementing effective planning after defining a specific goal. This problem has been the subject of extensive exploration across various domains. Several approaches necessitate the creation of a realistic simulator of the world [9], while others advocate the use of learned world models [8, 3]. However, these methods often demand a careful balance between precision and computational efficiency.

A critical facet of enabling robots to perform crow-like tool use is a deep understanding of affordance. The acquisition and application of affordance models represent a crucial juncture in this journey. Various techniques exist for learning and applying these models [5, 11], yet there remains a substantial distance to traverse before they can be flexibly employed across a diverse range of scenarios.

Effective actions in this domain can often be scarce and elusive, making the issue of insufficient exploration a persistent concern. This is also a central focus of the reinforcement learning community [4].

Moreover, realizing the transfer of knowledge between agents like crows introduces additional complexities. This endeavor might necessitate advancements in robot communication and imitation learning, enabling seamless knowledge sharing and collaboration.

In conclusion, the road to achieving crow-level tool use in robots is fraught with obstacles. Tackling these challenges will undoubtedly require a collaborative effort across various disciplines, pushing the boundaries of our understanding and capabilities in the realm of robotics, whether or not the proposed system in Sec. 2 is used.

References

[1] Jennifer Sootsman Buresh and Amanda L. Woodward. Infants track action goals within and across agents. *Cognition*, 104(2):287–314, August 2007. doi: 10.1016/j.cognition.2006.07.001. URL `https://doi.org/10.1016/j.cognition.2006.07.001`. 2

[2] György Gergely, Zoltán Nádasdy, Gergely Csibra, and Szilvia Bíró. Taking the intentional stance at 12 months of age. *Cognition*, 56(2):165–193, 1995. 2

[3] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023. 3

[4] Pawel Ladosz, Lilian Weng, Minwoo Kim, and Hyondong Oh. Exploration in deep reinforcement learning: A survey. *Information Fusion*, 85:1–22, 2022. 3

[5] Bogdan Moldovan, Plinio Moreno, Martijn Van Otterlo, José Santos-Victor, and Luc De Raedt. Learning relational affordance models for robots in multi-object manipulation tasks. In *2012 ieee international conference on robotics and automation*, pages 4373–4378. IEEE, 2012. 3

[6] Meiying Qin, Jake Brawer, and Brian Scassellati. Robot tool use: A survey. *Frontiers in Robotics and AI*, 9:1009488, 2023. 1, 2

[7] Christian Rutz, Lucas A. Bluff, Nicola Louise Reed, Jolyon Troscianko, Jason Newton, Richard Inger, Alex Kacelnik, and Stuart Bearhop. The ecological significance of tool use in new

caledonian crows. *Science*, 329(5998):1523–1526, September 2010. doi: 10.1126/science.1192053. URL https://doi.org/10.1126/science.1192053. 1

[8] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, Laurent Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy P. Lillicrap, and David Silver. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588(7839):604–609, December 2020. doi: 10.1038/s41586-020-03051-4. URL https://doi.org/10.1038/s41586-020-03051-4. 3

[9] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy P. Lillicrap, Karen Simonyan, and Demis Hassabis. A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*, 362(6419):1140–1144, December 2018. doi: 10.1126/science.aar6404. URL https://doi.org/10.1126/science.aar6404. 3

[10] Nat Geo WILD. Tool-making crows are even smarter than we thought | nat geo wild, January 2018. URL https://www.youtube.com/watch?v=UZM9GpLXepU. 1

[11] Natsuki Yamanobe, Weiwei Wan, Ixchel G. Ramirez-Alpizar, Damien Petit, Tokuo Tsuji, Shuichi Akizuki, Manabu Hashimoto, Kazuyuki Nagata, and Kensuke Harada. A brief review of affordance in robotic manipulation research. *Advanced Robotics*, 31(19–20):1086–1101, October 2017. doi: 10.1080/01691864.2017.1394912. URL https://doi.org/10.1080/01691864.2017.1394912. 3