



Robot tool use: A survey

OPEN ACCESS

EDITED BY

Chenguang Yang,
University of the West of England, United Kingdom

REVIEWED BY

Kai-Chieh Lin,
Tatung University, Taiwan
Frank Guerin,
University of Surrey, United Kingdom

*CORRESPONDENCE

Meiying Qin,
✉ meiying.qin@yale.edu,
✉ meiyingqin@gmail.com

SPECIALTY SECTION

This article was submitted to Robotic Control Systems, a section of the journal Frontiers in Robotics and AI

RECEIVED 05 September 2022

ACCEPTED 28 December 2022

PUBLISHED 16 January 2023

CITATION

Qin M, Brawer J and Scassellati B (2023).
Robot tool use: A survey.
Front. Robot. AI 9:1009488.
doi: 10.3389/frobt.2022.1009488

COPYRIGHT

© 2023 Qin, Brawer and Scassellati. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Meiying Qin*, Jake Brawer and Brian Scassellati

Yale Social Robotics Lab, Department of Computer Science, Yale University, New Haven, CT, United States

Using human tools can significantly benefit robots in many application domains. Such ability would allow robots to solve problems that they were unable to without tools. However, robot tool use is a challenging task. Tool use was initially considered to be the ability that distinguishes human beings from other animals. We identify three skills required for robot tool use: perception, manipulation, and high-level cognition skills. While both general manipulation tasks and tool use tasks require the same level of perception accuracy, there are unique manipulation and cognition challenges in robot tool use. In this survey, we first define robot tool use. The definition highlighted the skills required for robot tool use. The skills coincide with an affordance model which defined a three-way relation between actions, objects, and effects. We also compile a taxonomy of robot tool use with insights from animal tool use literature. Our definition and taxonomy lay a theoretical foundation for future robot tool use studies and also serve as practical guidelines for robot tool use applications. We first categorize tool use based on the context of the task. The contexts are highly similar for the same task (e.g., cutting) in *non-causal tool use*, while the contexts for *causal tool use* are diverse. We further categorize causal tool use based on the task complexity suggested in animal tool use studies into *single-manipulation tool use* and *multiple-manipulation tool use*. Single-manipulation tool use are sub-categorized based on tool features and prior experiences of tool use. This type of tool may be considered as building blocks of causal tool use. Multiple-manipulation tool use combines these building blocks in different ways. The different combinations categorize multiple-manipulation tool use. Moreover, we identify different skills required in each sub-type in the taxonomy. We then review previous studies on robot tool use based on the taxonomy and describe how the relations are learned in these studies. We conclude with a discussion of the current applications of robot tool use and open questions to address future robot tool use.

KEYWORDS

survey, robot tool use, human tools, manipulation, affordance

1 Introduction

Many robots are designed to interact with objects in the environment. Recent advances grant robots the ability to perform various tasks ranging from everyday tasks, such as swiping a card ([Sukhoy et al., 2012](#)), to professional tasks that require high precision, such as robot surgery ([Sarikaya et al., 2017](#)).

Among these tasks, robot tool use is gaining increasing attention. Being able to use human tools such as screwdrivers and scissors can greatly expand the applicability of a robot. Household robots will be able to assist humans better by performing a wider range of tasks with everyday tools; robots in chemistry labs will be able to run more experiments by leveraging the lab tools; manufacturing robots will be able to complete more tasks by utilizing construction tools without the need for specialized grippers. In this survey, a *tool* refers to the object attached to a robot. A *manipulandum* refers to the object being manipulated by the tool. An *object* is an umbrella term to include both tools and manipulanda.

Robot tool use requires three skills. The first skill is perception. A robot should identify and localize tools and manipulanda from the environment. For example, to drive a slotted screw, the robot needs to align the slotted screwdriver with the screw. Inaccurate pose perception of the screw will lead to misalignment, resulting in the failure of the tool use action. To successfully drive a screw, position knowledge alone is insufficient. In the above example, the tip of the screwdriver should both be at the position of the top of the screw, and oriented in a way that the flat tip of the screwdriver is aligned with the slot of the screw. Though challenging, the perception requirement is not unique to tool use. General robot manipulation also requires similar perceptual capabilities (Li et al., 2022).

The second skill of robot tool use is manipulation. Manipulation skills focus on how to realize the required kinematics and dynamics of tool use actions. The actions include two components as defined in Qin et al. (2021) and demonstrated in Figure 1: the contact poses and the course of the action. The contact poses include tool-manipulandum contact poses and gripper-tool contact poses (i.e., grasping). These poses consider both the translational (e.g., the tip of the pen should contact a point on the paper) and the rotational (e.g., the pen should contact the paper close to perpendicular to the plane of the paper, rather than parallel to it) relations of the tool and the manipulanda, or the tools and the gripper. Manipulation also encompasses both the tool trajectory (i.e., a time series of poses of a tool) and dynamics (i.e., the forces required for successful tool use). Though the manipulation skills required in tool use tasks may share similarity with general manipulation tasks, tool use additionally requires that a robot should update its body schema when a tool is held (Stoytchev, 2003). For example, general manipulation tasks consider exerting certain forces at the end-effector, while tool use tasks concern how the forces be generated at the tool rather than at the end-effector.

The third skill of robot tool use is high-level cognition. This includes reasoning and planning tool use actions given the tasks and available tools. For example, a robot may need to reason about how to grasp the tool to facilitate tool use, and determine the tool-manipulandum contact pose, the trajectory, and the force needed to use tools successfully. The robot also reasons about using a novel tool when learned tools are unavailable. Moreover, the robot should plan to use multiple tools to achieve a goal. Neurological evidence also supports that different cognitive processes were involved when a human uses a tool compared with separate hand and tool actions (Cabrera-Álvarez and Clayton, 2020).

The unique skills required by tool use distinguish tool use from general manipulation tasks (for a review on robot manipulation, see Mason, 2018; Kroemer et al., 2021). In the following sections, we start

with the discussion of what is tool use in robotics. Defining robot tool use is not a trivial task and we gain insights in animal studies which has a long history of studying tool use, especially how they distinguish tool use from general manipulations in animal behaviors. In Section 2, we present key results in animal studies and describe our definition of robot tool use. The definition not only sets a boundary for tool use, but also illuminates important skills of robot tool use. Moreover, researchers in animal studies also categorized different types of tool use and summarized the different skills required in each type of tool use. With this knowledge, we compile a tool use taxonomy in Section 3 and describe the skills required by different sub-types of tool use tasks in Section 4. The definition and taxonomy set a theoretical foundation for robot tool use, which is currently lacking in robotics research and acts. They also serve as a practical guideline for future applications. In Section 5, we organize robot tool use studies based on our taxonomy and focus on 1) the unique challenges of tool use tasks compared with general manipulation tasks in learning the required skills and 2) current advancements in how the skills are learned. We conclude this survey by discussing current applications of robot tool use and identifying the open challenges remaining in robot tool use.

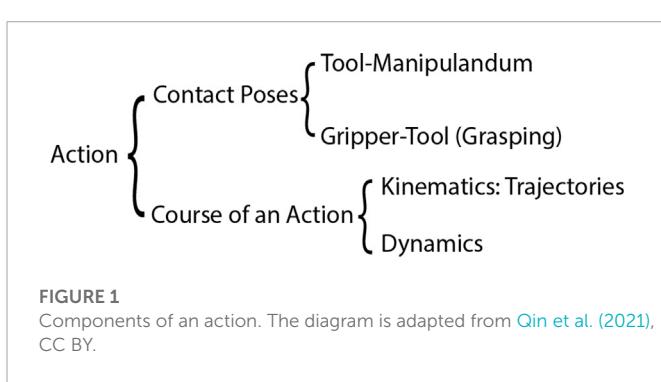
2 Definition of robot tool use

Defining robot tool use is necessary to understand the uniqueness of tool use as described in Section 1. Although the concept of robot tool use is intuitive, it is challenging to provide a precise definition to show clear distinctions from general manipulations. For example, one may consider robot tool use as robots using objects in the environment to achieve a goal. Under this definition, robots using refrigerators to store food should count as robot tool use. Along this line of logic, robots using the floor to support themselves should also be considered as robot tool use, which is counter-intuitive.

Researchers in animal tool use also encountered similar challenges. Tool use was initially considered a unique behavior only shown in humans (Oakley, 1944). With observations of tool use in animals, researchers debated which instances are genuinely tool use. For example, some may argue that a chimpanzee using a rock to crack a nut is tool use but cracking a nut against an anvil is not, while others may consider both cases are tool use. The debate urged a precise definition of animal tool use. In this section, we present key results in animal studies, gain insights from their argument, describe our definition of robot tool use, and explore the necessary aspects of tool use based on our definition. We focus on the implications of these animal studies for robot tool use, rather than on the implications for animal cognition. Therefore, the review of animal tool use studies is not meant to be comprehensive.

Van Lawick-Goodall (1970, p. 195) defined tool use as “the use of an external object as a functional extension of mouth or beak, hand or claw, in the attainment of an immediate goal,” emphasizing the goal-oriented and functional character of tool-use. Alcock (1972, p. 464) revised the definition by specifying the kind of objects that can be used as tools and identifying the scope of the goals: “Tool-using involves the manipulation of an inanimate object, not internally manufactured, with the effect of improving the animal’s efficiency in altering the form or position of some separate object.”

Beck (1980) identified a number of shortcomings with these definitions. First, only objects that are portable and manipulable should be considered as tools. Under this definition, dropping a



stone on an egg would be considered an example of tool use, but pounding a fruit on a tree would not be. The latter case is considered proto-tool-use (Parker and Gibson, 1977). Second, an agent should understand the connection between the goal and the tool. Otherwise, the conditioned behavior of a rat pressing a lever in a Skinner box would be considered tool use, and Beck considered this inappropriate. Third, the tool need not be externally manufactured to the agent using it nor inanimate. Researchers observed that captive apes threw feces toward human intruders, and a chimpanzee utilized the dead body of a colobus monkey to hit a conspecific, suggesting that a live ape could be utilized in a similar way. Beck argued that these behaviors should be considered tool use. Fourth, the goal of tool use can be extended beyond feeding or drinking to other goals such as self-maintenance. As a result, Beck (1980, p. 10) re-defined tool use as “the external employment of an unattached environmental object to alter more efficiently the form, position, or condition of another object, another organism, or the user itself, when the user holds or carries the tool during or just prior to use and is responsible for the proper and effective orientation of the tool.”

For decades, Beck’s definition has been accepted widely in the field of animal cognition and was even adopted in early robot tool use studies (e.g., Stoytchev, 2007). Two observations motivated St. Amant and Horton (2008) to propose a new definition of tool use. Krützen et al. (2005) reported that dolphins hold marine sponges in their rostrum in order to prevent potential injuries when probing for food. Breuer et al. (2005) observed that a wild gorilla tested the depth of water with a stick while it walked across a pond. These two behaviors fall outside of Beck’s definition of tool use since they do not involve altering the state of another object. St. Amant and Horton were also concerned about Beck’s definition that it over-emphasized peripheral aspects of tool use, such as the unattached property; an animal can use a stick that is still attached to a tree as a tool. Moreover, they argued that Beck’s definition is vague to determine whether a goal was achieved accidentally. They observed that purposeful behaviors require a continuum of control. Therefore, St. Amant and Horton (2008, p. 1203) re-defined tool use as “the exertion of control over a freely manipulable external object (the tool) with the goal of 1) altering the physical properties of another object, substance, surface or medium (the target, which may be the tool user or another organism) via a dynamic mechanical interaction, or 2) mediating the flow of information between the tool user and the environment or other organisms in the environment.” They elaborated that the interactions between tools and manipulanda should be dynamic. Under this definition, stacking boxes to reach bananas is not tool use since the interactions between boxes remains fixed once they have been stacked, while cracking a nut with rock is tool use because the interactions between the nut and the rock is constantly changing.

Shumaker and Walkup joined Beck to revise the Beck’s widely accepted definition and incorporated St. Amant and Horton’s argument: “the external employment of an unattached or manipulable attached environmental object to alter more efficiently the form, position, or condition of another object, another organism, or the user itself, when the user holds and directly manipulates the tool during or prior to use and is responsible for the proper and effective orientation of the tool.” (Shumaker et al., 2011, p. 36) We based our definition of robot tool use on this revised definition.

Other definitions of tool use in animal studies exist. Some may simply be shorter versions of these definitions (e.g., Chevalier-Skolnikoff, 1989; Matsuzawa, 1999). Others may disagree with the

scope of tool use. For example, Asano (1994) did not restrict the tools to be something being held. This might result in the scope of tool use being overly broad since any behavior may eventually count as tool use, such as walking, which utilizes the ground as the “tool”. Lestel and Grundmann (1999) expands the scope of tool use even more by including abstract concepts such as culture as potential tools. These discussions may be too philosophical and lack operational details for robotics research.

We identify three essential points in these definitions. First, tool use must have a goal, despite a lack of consensus regarding a goal’s scope. Second, instead of achieving a goal through random exploration, an agent utilizing a tool should understand the connection between the goal and the behavior. Third, the tool should satisfy specific physical criteria, such as being freely manipulable. Based on these points, we define *robot tool use* as:

A robot attaches or secures to its end-effector an external, unanimated, freely available object or an object attached to another object, in order to achieve a goal of altering the state of another object, updating its own state, or other goals, through purposeful manipulations.

Our definition adopts Shumaker et al.’s definition with minor modifications. First, we restrict the tools to be externally manufactured and unanimated. Unlike living creatures, a robot typically does not produce materials (e.g., feces, spider webs) from its body. We require the tools to be unanimated because an animated object that a robot would most likely manipulate is another robot. We consider this a better fit to the area of multi-agent systems rather than tool use since it involves synchronization and communication between robots. Second, we relax the interactions between the tool and object to be manipulated to be dynamic or static. Therefore, using a container to relocate other objects would count as tool use. Third, we relax the goal of tool use. The scope of the goals in animal tool use was summarized based on animal behavioral observations. Given that tool use in animals is structurally simple even in non-human primates (Fragaszy and Eshchar, 2017), the goals in the above definitions are restricted to altering the state of another object and updating one’s knowledge about the environment. In contrast, as robots often utilize human tools, robot tool use is motivated by the same goals for which these tools were designed, goals that can far exceed in scope and complexity those observed in animal studies. On the contrary, robots are required to utilize human tools. The design of human tools is more complex than those used by animals so that human tools may serve purposes beyond the goals identified in animal studies. Therefore, we prefer not to restrict the scope of the goal of robot tool use.

3 A taxonomy of robot tool use

Researchers of animal tool use recognizes that there are different types of tool use. For example, antlions mechanically throwing sand to capture preys in the same manner across all contexts is notably different from chimpanzees carefully adjusting sticks to fish termites even in the same context. Similarly, robot tool use also has many different types. A robot being pre-programmed to cut pizza with a particular knife is very different from it learning to adjust the gestures when presented with different knives. In this section, we overview

taxonomies proposed in animal and robotics studies, and present a taxonomy on robot tool use.

3.1 Taxonomy of animal tool use

[Alcock \(1972\)](#) proposed a dichotomy of tool use in animals: *stereotyped tool use* is seen mostly in invertebrates and fish and *flexible tool use* is typically seen in birds and mammals. [Hunt et al. \(2013\)](#) considered this dichotomy an accurate description of two fundamental types of tool use with different underlying processes, despite being oversimplified. Stereotyped tool use is inherited and animals only utilize tools in default ways in particular contexts. Examples of stereotyped tool use include antlions throwing sand to capture preys ([Alcock, 1972](#)). The species of antlions developed this behavior from the pre-existing non-tool use behavior of random sand flicking in order to maintain their pits. The tool use behavior of antlions throwing sand evolves as a phenotypic change in this species. As a result, these behaviors are widespread across the species of antlions and rarely vary within and across individuals.

In contrast, flexible tool use, which is also referred to as creative tool use, is learning-based that animals explicitly reasoning about the usages based on the context. It is this type of tool use that some believe signals intelligence ([Call, 2013](#)) and interests researchers in animal cognition. Chimpanzees' cracking nuts with rocks and fishing termites with sticks are examples of flexible tool use ([Biro et al., 2003; Lonsdorf, 2006](#)), as a juvenile chimpanzee acquires such skills by observing its parent(s). Therefore, the learning happens at the level of the individual, rather than at the level of genus. Indeed, each instance of nut cracking or fishing termites can be differentiated even within the same context by the same chimpanzee. Unlike stereotyped tool use, flexible tool use does not share context-dependency and thus can occur across different contexts. We summarized the differences between stereotyped tool use and flexible tool use suggested by Hunt et al. in [Table 1](#).

[Call \(2013\)](#) further identified different types of flexible tool use from the perspective of problem-solving in terms of creativity and adaptivity.

- Solving novel problems with old solutions;
- Solving old problems with novel solutions;
- Solving novel problems with novel solutions.

Solutions may include utilizing one tool, selecting a tool from available options, manufacturing a novel tool, or using multiple tools sequentially.

In contrast to the above taxonomies, [Wimpenny et al. \(2009\)](#) categorized tool use based on the number of tools involved in a problem but overlooked the complexity of the decision process. [Boesch \(2013\)](#) categorized tool use based on four levels of increasing complexity though in some sense reminiscent of Wimpenny et al.'s approach.

- *Simple tool use*: Using one tool, e.g., a chimpanzee uses a twig for fishing termites ([Goodall, 1964](#)). The animal only needs to understand the connection between itself and the reward *via* the tool, which is a first-order problem ([Visalberghi and Fraszy, 2006](#));
- *Combined tool use*: Using two tools simultaneously, e.g., a capuchin monkey uses a rock to pound a nut on a hard

surface ([Spagnoletti et al., 2011](#)). The animal needs to consider both spatial relationships concurrently to connect itself with the reward, which is a second-order problem ([Visalberghi and Fraszy, 2006](#));

- *Sequential tool use*: Using multiple tools one after another, including using a tool for manufacturing another tool, e.g., a chimpanzee using multiple tools in sequence to break a bee hive, open honey chambers, and extract the honey. This behavior not only requires the animal to keep in mind multiple causal relationships sequentially and choose the correct sequence, but also imposes temporal delay for the reward;
- *Composite tool use*: Combining multiple tools to use as one tool, a tool use behavior yet to be discovered in animals and currently unique to humans.

3.2 Taxonomy of robot tool use

While the above taxonomies are based on animal studies, [Tee et al. \(2018, 2022\)](#) proposed a categorization based on default usages of tools in robotics, and identified three types of tools: category-I tools that "help to amplify/augment certain kinematic or dynamic aspects of functions that are already in an agents repertoire." (p. 6439), category-II tools that are similar to category-I tools but "require actions different from what the agent would have performed, without the tool, to achieve these functions." (p. 6439), and category-III "provide new functions that a human cannot perform without a tool." (p. 6440) As an example, they categorized a vacuum cleaner as a Category-III tool because a robot cannot perform a cleaning task without this tool. However, a vacuum cleaner can be used as a rake to reach objects or as a hammer to hit objects in other contexts. In these contexts, the vacuum cleaner should be classified as category-I tools. Given that this categorization does not consider contexts of tool use, it will be challenging for a system following this categorization to perform flexible tool use, which is context-based.

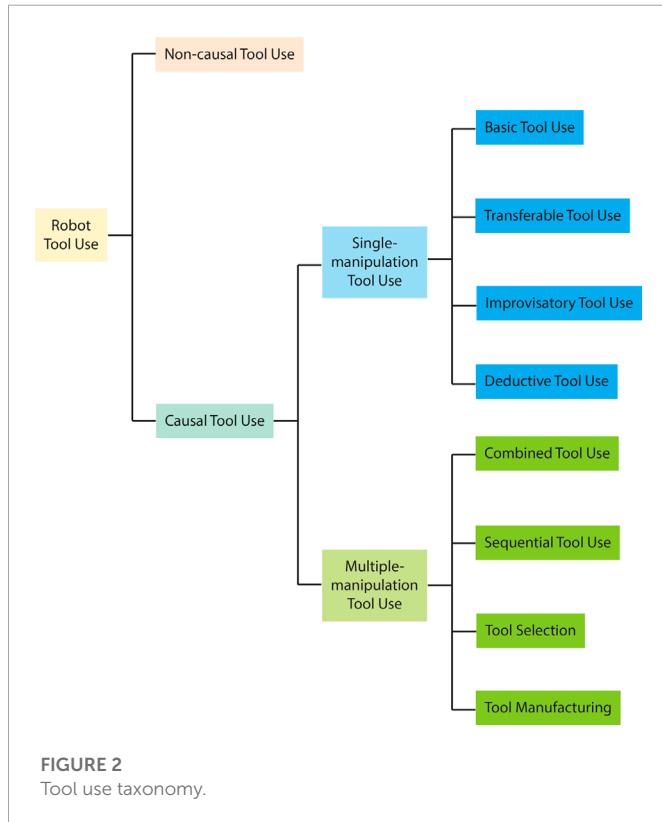
Based on the taxonomies of animal tool use and the characteristics of robot tool use, we devise a taxonomy as shown in [Figure 2](#). We categorize robot tool use into *non-causal tool use* and *causal tool use*, which are similar to stereotyped tool use and flexible tool use in animals, respectively. We changed the terminology for two reasons. First, we would like to emphasize the fundamental differences between the two types of tool use behaviors in robots regarding whether robots should understand required causal relations, which are elaborated in [Section 4](#). Second, though we consider it necessary for a robot to understand required causal relations in order to achieve behaviors similar to flexible tool use, there is a lack of evidence showing the mechanism of flexible tool use in animals. Therefore, we would like to avoid claiming that flexible tool use in animals is causal-based, and such discussion is beyond the scope of this survey.

We further categorize causal tool use into *single-manipulation tool use* and *multiple-manipulation tool use* based on Boesch's taxonomy. A single manipulation refers to being presented with a single tool and using the tool to perform one action (e.g., pushing, scooping) in order to achieve one goal, though a robot may observe the usage of multiple tools to learn a task. Multiple manipulations may involve one or any combinations of multiple tools, multiple actions, and goals consisting of multiple sub-goals.

Inspired by Call's taxonomy, we categorize single-manipulation tool use into *basic tool use*, *transferable tool use*, *improvisatory tool use*,

TABLE 1 Comparison of stereotyped tool use and flexible tool use based on Hunt et al. (2013)'s descriptions.

	Stereotyped tool use	Flexible tool use
Distribution	genus level	individual level
Development	phenotypic changes stems from pre-existing non-tool use behaviors	observational learning
Variability	almost no variations within and between individuals	very different within and between individuals



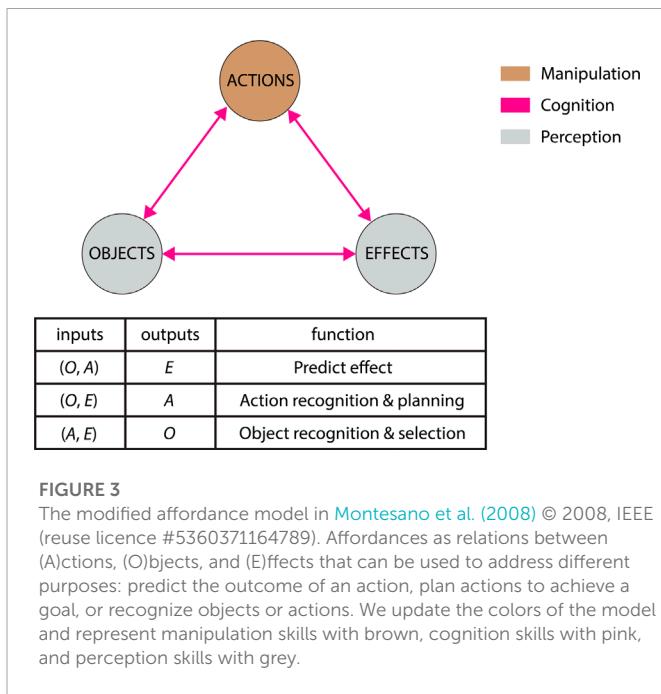
and *deductive tool use*. Basic tool use is the most basic form of tool use. In basic tool use, a robot uses a learned tool to solve a learned task, such as pushing a block, striking a xylophone, and cutting a cake. Unlike non-causal tool use that exclusively focuses on the actions, basic tool use focuses on the causal relations between actions and effects. Transferable tool use is a more complicated form of tool use, which aims to transfer learned tool use skills to other intra-category objects that share common form factors (e.g., using mugs of different shapes to pour liquid into different containers). Improvisatory tool use adds further complexity by generalizing learned tool use skills to novel inter-category objects. These objects are generally not designed to perform these tasks, such as using the handle of a screwdriver in place of a hammer to drive a nail. Deductive tool use concerns the problem of using a novel tool to solve a novel task. A robot will not be provided any prior knowledge about the tool or the task, nor given opportunities to learn about them from demonstrations. Instead, the robot should deduct the usage of a tool from its physical knowledge about the world.

We categorize multiple-manipulation tool use into *combined tool use*, *sequential tool use*, *tool selection*, and *tool manufacturing*.

Combined tool use and sequential tool use are similar to the definitions as in Boesch's taxonomy, though sequential tool use does not include constructing a new tool in our definition as it requires more sophisticated manipulation skills. Tool selection refers to the process of choosing the most appropriate tool among many options in order to complete a tool use task. Shumaker et al. (2011) defined tool manufacturing as "simply any structural modification of an object or an existing tool so that the object serves, or serves more effectively, as a tool." This definition only includes modifying an existing object. We combine this definition with composite tool use in Boesch's taxonomy, and re-define tool manufacturing as **the process of modifying or combining objects or existing tools, with or without the usage of other tools, to complete a tool use task, or to complete the task more efficiently**. Different from Shumaker et al. and Boesch's definition, our definition also explicitly includes the possibility of utilizing other tools in the process of manufacturing.

We do not enforce a subdivision of multiple-manipulation tool use by difficulty, unlike a comparable category in Boesch's taxonomy. Boesch was able to rank the categories in animal tool use because the types of tools leveraged by non-human animals are comparatively limited, and the manipulation skills in these animals are usually relatively simple. In robot tool use, the difficulty is dependent on the actual problem to solve, rather than the category that the problem belongs. For example, utilizing two tools in sequence may be simpler than creating a new tool that requires sophisticated manipulation skills. However, a problem that requires planning to use ten tools may be more challenging than a problem that requires a robot to combine two parts as a new tool.

As a summary, the criterion of the top-level classification is whether the context exhibit much variations across instances of tool use in the same task. As a result, robot tool use is classified into non-causal tool use and causal tool use. Causal tool use is further categorized into single-manipulation tool use and multiple-manipulation tool use based on the complexity of the task, which is quantified by the number of actions, goals, and tools. We categorize single-manipulation tool use into four sub-types. Three of the sub-types (i.e., basic tool use, transferable tool use, and improvisatory tool use) rely on prior experiences of tool use and the other (i.e., deductive tool use) only requires knowledge of physical rules or experiences of general manipulation. The three sub-types differ from each other in terms of object features. Single-manipulation tool use can be considered as building blocks of causal tool use, and multiple-manipulation tool use combines the building blocks in different ways. The different combinations form the sub-types of multiple-manipulation tool use, which are combined tool use, sequential tool use, tool selection, and tool manufacturing.



4 Required skills of robot tool use

Our definition of tool use has three important components: objects, goals or desired effects, and manipulations or actions to achieve the goals. These three components agree with the three ingredients of the affordance model defined by Montesano et al. (2008). The affordance model attempts to provide an operational definition of the concept of affordances, whose precise definition is still debatable (for a review, see Jamone et al., 2016; Zech et al., 2017). The concept was first introduced by Gibson (1979, p. 127) as what the environment “offers the animal, what it provides or furnishes, either for good or ill”. Despite the lack of consensus around its definition, the concept of affordances has facilitated much robotic research (Stoytchev, 2005b; Cakmak et al., 2007; Moldovan et al., 2012, 2013; Katz et al., 2014; Ruiz and Mayol-Cuevas, 2018; Lueddecke et al., 2019).

The affordance model by Montesano et al. formulated affordance as three-way relations between objects, actions, and effects. Figure 3 shows our modified version of this affordance model with coloring. The coloring captures the three skills required for robot tool use as described in Section 1. While generating actions requires manipulation skills, perceiving the effects and the objects demands perception skills. Understanding the connections between the nodes in the model needs cognition skills. The key to robot tool use is to understand tool affordances.

Each subtype of tool use in our taxonomy addresses different aspects of affordances and requires different skills, as shown in Figure 4. Non-causal tool use focuses on generating desired motions, which correspond to the action node in the affordance model. While the manipulation skills are similar to the ones in general manipulation tasks, tool use tasks require additional manipulation skills, such as updating a robot’s body schema when a tool is attached to its gripper. This type of tool use requires a robot to focus on the actions, though without the need to consider the objects, the effects, or the relations.

Causal tool use involves learning and applying tool affordances, which focuses on cognition skills. Single-manipulation tool use learns and reasons about affordances. Among the different types of single-manipulation tool use, basic tool use learns how to achieve desired effects with actions. With learned relations between actions and effects, transferable tool use learns the relations between tools and actions in order to adjust actions based on novel tools that share similar form factors with the learned tools. In addition to these two relations, improvisatory tool use requires a robot to understand what specific tool features cause the effects so that it can generalize learned skills to inter-category objects. As a result, improvisatory tool use requires a robot to learn the entire affordance model. These tool use tasks generalize tool use to novel objects by learning and inducing affordance from observations. In contrast, deductive tool use requires a robot to complete a tool use task without prior knowledge using unlearned tools. As a result, the robot has no information to induce affordances and should perform deductive reasoning from general physical rules. In short, basic tool use, transferable tool use, and improvisatory tool use requires the incremental learning of more relations from tool use demonstrations, while deductive tool use requires a robot to be able to infer rather than learn the relations from demonstrations of general manipulation tasks or physical rules provided directly.

The challenge for multiple-manipulation tool use is to apply the affordances effectively rather than learning the affordances. Tool manufacturing requires more sophisticated manipulation skills; sequential tool use and tool selection require higher cognition skills; tool manufacturing requires a high level of manipulation and cognition skills. We summarize the additional skills required in each sub-type in Table 2, and we elaborate more in Section 5.3.

As a summary, while non-causal tool use focuses on characterizing and replicating actions, causal tool use considers the relations between actions and other skills in addition. Under causal tool use category, single-manipulation tool use focuses on learning the skills and the relations while multiple-manipulation applies this knowledge to solve more complicated tasks and may require more sophisticated skills.

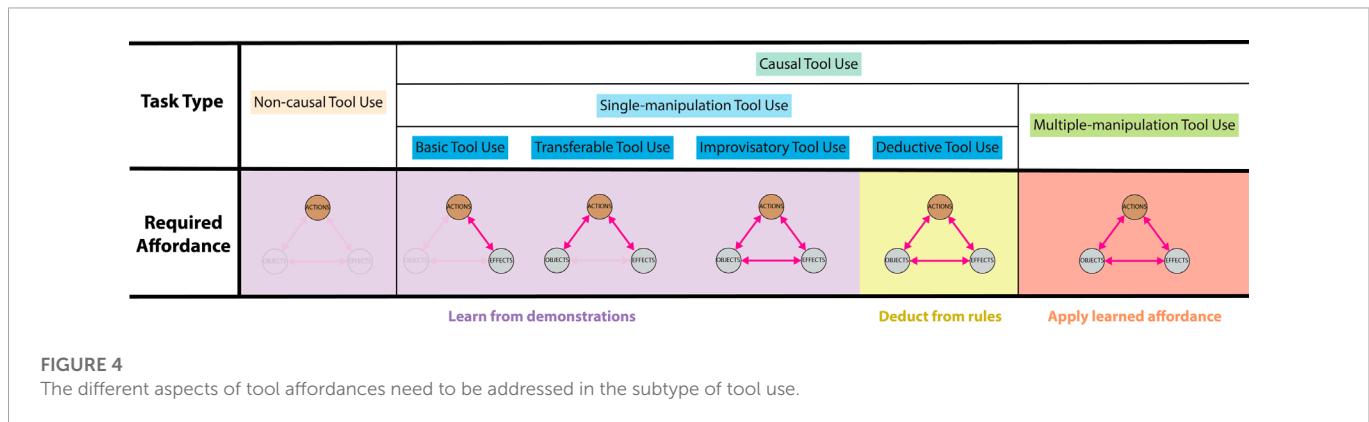
5 Robot tool use literature

In the previous sections, we present our definition of robot tool use drawing from the animal studies literature. Our definition highlights the three important skills in tool use, echoing the three-way relations between actions, objects, and effects in standard affordance models. We also describe a taxonomy of tool use with each type of tool use requiring different skills or relations.

In this section, we organize robot tool use literature based on our taxonomy. As different levels of assumptions can be made, we would like to point out that techniques that focus on the unique challenge in a higher-level tool use may not necessarily allow a robot to handle the challenges that belong to lower-level tool use. In each section, we focus on the techniques that can solve the unique challenge at each level of tool use. We would also like to emphasize that the purpose of robot tool use is not to mimic or model how animals use tools, but rather to allow robots to use tools.

TABLE 2 Additional skills required in multiple-manipulation tool use beyond single-manipulation tool use.

Additional manipulation skills		Additional cognition skills	
		Reasoning	Planning
Combined Tool Use	Coordinate the use of multiple tools	Choose appropriate parameters for each tool use	N/A
Sequential Tool Use	N/A	N/A	Plan the order of using multiple tools
Tool Selection	N/A	Choose appropriate tools among many available tools	N/A
Tool Manufacturing	Perform the actions of assembling different parts together as a tool, or modify the current tool	Choose appropriate parts to be assembled and decide where to attach each part, or decide the desired state of the unmodified tool	Planning the order of the manufacturing, which may include sequential tool use



5.1 Non-causal tool use

Non-causal tool use is the ability to use learned tools to solve learned tasks, without understanding the cause-and-effect relationship between the actions and the goals. The purpose of non-causal tool use is to duplicate or reproduce actions with limited variations. It could be achieved by programming a wide variety of tool use actions such as nut fastening in aircraft production that requires high precision (Pfeiffer et al., 2017), stub grinding and deburring with force control (Robertsson et al., 2006), handwriting that involves multi-contact manipulation (Kim et al., 2014), furniture polishing that uses an impedance model (Nagata et al., 2001), generating a collision-free polishing path (Takeuchi et al., 1993), accurately drawing a circle with a compass that involves complex contacts (Kutsuzawa et al., 2017), unfastening screws in collaborative tasks (Li et al., 2020), and pouring based on the volume of liquid (Rozo et al., 2013), or actions relevant to tool use such as grasping a knife resting on a cutting board that requires a high level of dexterity (Xue and Jia, 2020), or segmenting a surgical tool from the background while using it (Garcia-Peraza-Herrera et al., 2017; Su et al., 2018). The purpose of these approaches is to automate one process to facilitate human work. Therefore, the implementations are designed to be highly specific to the task.

However, given the wide range of tasks, it is impractical to program all tool use tasks. Being able to learn these tasks is desired. One approach is to treat tool use tasks the same way as general manipulation tasks and learn the actions accordingly. One of the classic algorithms of learning actions is dynamic movement primitives (DMP) (Schaal, 2006; Ijspeert et al., 2013). DMP leverages the concept of attractors from dynamical systems, and actions are represented as

a set of linear differential equations. A more intuitive approach to understanding DMP is to visualize the equations as vector fields, where a trajectory is formed by following the vectors from a starting point to an end point. Each dimension may need to be learned separately and then coupled together. One advantage of DMP is that the shape of the trajectory can be distorted based on the starting point and the end point. DMP and its variations have been demonstrated with tool use tasks such as swinging a tennis racket (Ijspeert et al., 2002; Schaal, 2006), playing table tennis (Muelling et al., 2010), playing ball-in-a-cup (Kober et al., 2008), pouring liquid (Pastor et al., 2009), and whiteboard cleaning (Kormushev et al., 2011). Algorithms other than DMP have also been employed to represent action primitives, such as probabilistic movement primitives (Paraschos et al., 2013) and Fourier movement primitives (Kulak et al., 2020). Actions were also parameterized as minimal plans to facilitate action interpretation (Guha et al., 2013). To handle tasks that require high manipulation precision such as using chopsticks, model-free imitation learning was chosen (Ke et al., 2021). For tasks that do not require high precision of force or position control, indirect force controllers (Lutscher and Cheng, 2013) or a unified algorithm for dynamic object manipulation (Tsuiji et al., 2015) can be considered. While these are methods designed to learn actions, general learning methods such as deep learning (Droniou et al., 2014; Byravan and Fox, 2017) and reinforcement learning (Peters and Schaal, 2006) were also used.

While these studies focus on learning actions, others focus on segmenting continuous actions into action primitives for tool use (Ramirez-Amaro et al., 2014a; Lioutikov et al., 2017). A similar line of research on general manipulation tasks is to recognize the tasks based on the classification of actions (Ramirez-Amaro et al., 2014b; 2015;

[Hu et al., 2014](#); [Wölfel and Henrich, 2018](#); [Shao et al., 2021](#); [Koch et al., 2022](#)). This approach attempts to ground action profiles to labels, either primitive labels or task labels, and do not relate actions to the effects on the objects being manipulated. For example, the whiteboard swiping action will be characterized as the translational movement of the eraser in this approach, rather than the words being erased, which is the effect. While grounding action profiles to labels is useful in some applications, it does not permit causal tool use.

The above studies treated tool use tasks in the same manner as general manipulation tasks. As a result, they cannot adjust actions based on how the tools are grasped since they do not have tool-related knowledge. In order to accommodate the tools attached to the end-effector, a robot needs to update its body schema to include the tools, or in other words, to calibrate the tool in the gripper. Prior studies focus on updating robot kinematics by considering the tip of a tool [e.g., [Kemp and Edsinger \(2006\)](#)], which is considered the primary contact point between the tool and the environment. Among these studies, some manipulated the tool with kinematic control ([Stoytchev, 2003](#); [Nabeshima et al., 2005](#)), and others found it necessary to perform dynamic control ([Kemp and Edsinger, 2006](#); [Nabeshima et al., 2007](#); [Jamone et al., 2013](#); [Hoffmann et al., 2014](#); [Karayannidis et al., 2014](#)). Despite these studies' success, considering the tool's tip only is insufficient for all tool use tasks. For example, it is insufficient to know where the tip of a mug is when it is used to pour liquid into another container. The mug needs to be tracked with multiple markers attached to it ([Lee et al., 2008](#)). Another example is joint tools such as a pair of scissors. In this scenario, a grounded relational representation of the entire tool is needed ([Katz et al., 2008](#)). Beyond tool calibration, other studies explored collision detection ([Colgate et al., 1995](#)) and obstacle avoidance ([Lee and Song, 2021](#)) with tools attached to the gripper, as well as robot motion planning to complete tool use tasks ([Kobayashi and Hosoe, 2009](#); [Holladay et al., 2019](#)) or planning for the grasping of the tool ([Lin and Sun, 2015](#); [Chen et al., 2019](#); [Raessa et al., 2019](#)) in robot motion generation.

5.2 Causal tool use—Single-manipulation tool use

5.2.1 Basic tool use

Though both basic tool use and non-causal tool use leverage learned tools to solve learned tasks, basic tool use can adjust actions based on the desired effects while non-causal tool use cannot. In other words, basic tool use requires robots to understand the causal relations between actions and effects. For example, a robot performing basic tool use is able to push an object further away given a target region that is further away, while a robot performing non-causal tool use will simply attempt to duplicate learned actions and does not adjust the actions based on the target region that is further away.

[Sinapov and Stoytchev \(2008\)](#) conducted an early study to explore the relation between actions and effects with motion babbling. They utilized six different tools (T-stick, L-stick, straight stick, L-hook, Y-hook, and an arrow-shaped tool) to relocate a puck with six pre-defined exploratory behaviors (i.e., push, pull, slide-left, slide-right, rotate-left, and rotate-right). For each tool, the robot learned the distribution of movement trajectories of the puck relative to its starting location. [Forestier and Oudeyer \(2016\)](#) employed an active version of Model Babbling to explore the distribution of manipulanda after tool use with two different sticks. These studies focus on the

potential distribution of the location of manipulanda, rather than the one-to-one relationship between an action and its effect. Therefore, it is challenging to utilize tools to achieve desired effects with this method.

Other studies learned the one-to-one relation of an action and its effect, though in a quantitative manner. [Okada et al. \(2006\)](#) focused on verifying the effects as success or failure of tool use tasks such as pouring. [Pastor et al. \(2011\)](#) focused on predicting whether an object has been successfully struck by a pool cue or flipped using chopsticks. Studying the relation of an action and its effect in this manner is suitable if the state of the effects is discrete, but may not fit tool use tasks whose effects are continuous such as pushing an object 10 cm to its right.

Studies that focus on learning the one-to-one relation of an action and its effect in a qualitative way generally employed tasks that result in the relocation of manipulanda. [Stoytchev \(2005a, 2008\)](#) pre-defined eight pulling actions and recorded the effects of these actions with five different tools into an affordance table. As the actions were discretized, the effects can also be categorized in discretized space. In the evaluation, a robot needed to choose appropriate actions based on the affordance table in order to pull the manipulanda into a goal region, given one of the learned tools. Though [Tikhanoff et al. \(2013\)](#) also leveraged pre-defined actions, they allowed the actions to be parameterized with continuous variables, e.g., a randomly sampled pushing direction. Rather than keeping an affordance table, they used Least Square Support Vector Machines to regress the actions to the effects. [Elliott et al. \(2016\)](#) considered more types of push and pull. They also leveraged two regression techniques: linear regression and Gaussian process regression. Other than pulling and pushing tasks, [Elliott and Cakmak \(2018\)](#) explored cleaning tasks to relocate dirt. As the manipulanda are clusters of rigid bodies rather than a single rigid-body manipulandum, they represented the surface as a grid, and trained a pixel-level classifier to predict whether each pixel contains dirt after an action. The robots in the above studies explored tool use by themselves, pre-defined actions are necessary. In contrast, [Liu et al. \(2018\)](#) did not pre-define actions and took the method of imitation learning and learned with deep reinforcement learning.

These studies focus on pushing and pulling tasks. A common feature of these tasks is that the desired effect determines how a tool should contact a manipulandum. Other tool use tasks may permit multiple equally viable ways for a tool to make contact with a manipulandum to achieve the same effect. For example, pouring liquid from different orientations all result in the same effect of a container being filled. [Claassens and Demiris \(2011\)](#) conducted preliminary studies and termed such properties with affordance symmetries. Affordance symmetries are important because a robot will be able to generate different trajectories to complete a tool use task when the learned contact results in collision. However, few studies have explored this direction to our knowledge.

5.2.2 Transferable tool use

Transferable tool use describes the ability to take tool use skills trained on an object to other intra-category objects defined by a common form factor. Therefore, the key to transferable tool use is to match the unlearned objects with learned objects.

We first present studies that concern specific types of tool use tasks. Most of these focused on relocation tasks such as pulling and pushing. [Mar et al. \(2017\)](#) and [Nishide et al. \(2011\)](#) leveraged self-organized maps to extract tool features to avoid the need to pre-defining the features. [Takahashi et al. \(2017\)](#) learned a model with a

deep neural network that incorporated both grasping information and tool functions. [Vogel et al. \(2017\)](#) searched for the “sweet spot” of a novel baseball bat-like object when used to hit a baseball by sensing the force at the end-effector. Other studies considered pouring tasks. [Kroemer et al. \(2012\)](#) used a kernel-based approach to generalize learned action skills to novel objects. [Brandi et al. \(2014\)](#) performed warping to the point cloud of a learned container to match a novel container. [Dong et al. \(2019\)](#) adjusted pouring behavior by estimating the volume of liquid in the unknown containers. While these studies attempted to transfer tool use skills to novel tools, other studies explored how to act upon novel manipulanda. [Gemici and Saxena \(2014\)](#) sought to transfer cutting skills to food of varying physical properties such as hardness. [Elliott et al. \(2017\)](#) transferred learned surface cleaning actions to different surfaces, including surfaces of different sizes. [Li et al. \(2018\)](#) developed the Push-Net so that the system can push novel objects for re-positioning and re-orientation.

Though these studies demonstrated promising results on specific tool use tasks, it is unknown whether these algorithms could generalize to other types of tool use tasks. Therefore, other researchers investigated algorithms that transfer learned skills more broadly and demonstrated with multiple tool use tasks. [Tee et al. \(2018, 2022\)](#) matched the point cloud of unseen tools to the point cloud of the end-effector and arms of the robot to obtain the usage of the tools. The kPAM/kPAM 2.0 ([Manuelli et al., 2019; Gao and Tedrake, 2021](#)) used keypoints on the tools to represent shared global shape of the category of tools, and tool use skills were inferred from these keypoints. [Stückler and Behnke \(2014b\); Stückler et al. \(2016, 2013\); Stückler and Behnke \(2014a, 2015\)](#) considered the point cloud presentation of tools and performed deformable registration with different levels of resolution in order to match the overall shape of the tools.

The approach of these studies requires two steps: one to learn tool use skills as basic tool use, and one to learn the transfer process. Other studies merged the two steps and learned them in one step. [Sinapov and Stoytchev \(2007\)](#) incorporated the shape of the tool when learning tool use models for the pulling task. As a preliminary model, transfer was only demonstrated with tools of the same shape but different sizes. [Gonçalves et al. \(2014a,b\)](#) utilized a Bayesian network to learn how the actions and tool shapes influence the effects. The shape parameters include area, convexity, eccentricity, compactness, circularness, and squareness. Due to the large size of the network, it needed to be reduced to be able to train effectively. They validated their technique with pulling and pushing tasks. [Dehban et al. \(2016\)](#) took a similar approach but overcame the drawback of the need for a discretization of data. To be able to handle grasping, [Mar et al. \(2015\)](#) leveraged support vector machines to map geometric features between learned and novel tools for pulling.

There are pros and cons of these two approaches. Training everything in one step may be more convenient, but the feature space can be quite large and requires more data. Training in a modular way will make it easier to diagnose when the algorithm does not function as intended. It will also make it easier to modify or incorporate new features as the former requires the entire model to be retrained.

5.2.3 Improvisatory tool use

Improvisatory tool use describes the ability to use tools in a creative way, which involves generalizing learned tool use skills from objects designed for the tasks to inter-category objects. These objects may not share common form factors with the canonical tools.

Therefore, local features of the tools that lead to the desired effects should be identified.

While transferring tool use requires a robot to infer how actions are affected by novel tools given the relation between actions and effects, improvisatory tool use requires a robot to also understand what features of the tools caused the effects, which is the relation between tools and effects. In other words, improvisatory tool use calls for the learning of the full affordance model ([Montesano et al., 2007; 2008](#)).

In order to identify local features in unlearned tools, the function of a tool needs to be detected on a per-part basis. Insights can be gained from a related line of research that explores task-oriented grasping of novel objects. These studies made efforts to detect the functional part of a tool in different tasks so that the system can generate different grasping of the same object based on the task ([Myers et al., 2015; Song et al., 2010; 2011b; a; Ek et al., 2010; Madry et al., 2012; Song et al., 2015; Murali et al., 2020; Kocic et al., 2017; Detry et al., 2017](#)). Similar to these studies, studies that focus on part detection for tool use also leveraged geometric features. [Schoeler and Wörgötter \(2015\)](#) segmented the tools and used graphs to represent the relations between different tools parts. [Nakamura and Nagai \(2010\)](#) learned the full affordance model. They provided human static demonstrations without showing the course of actions, and detected local features with the Scale Invariant Feature Transform.

Given the functions of each tool part alone, a robot cannot realize improvisatory tool use since a robot have no knowledge about how to orient a tool. The robot needs to combine the tool parts information with tool use knowledge. Due to challenges in modeling grasping, [Fitzgerald et al. \(2019\)](#) achieved the goal with human-guided adaptation that gained information on how to improvise each tool from human demonstrators. To improvise tool use without the need of human demonstrations for each tool, [Agostini et al. \(2015\)](#) learned actions with a modified DMP and used a Repository of Objects and Attributes with Roles to detect potential usages of a tool. This method is based on matching the global shapes of tools. Though this method can perform some improvisatory tool use (e.g., utilizing a knife vertically for stirring in a way similar to using a spatula), the transfer is limited. Other studies considered both global and local features. [Fang et al. \(2020\)](#) and [Xie et al. \(2019\)](#) took 2D images as input and trained neural networks for improvisatory tool use. While these studies learned tool use skills and tool feature detection together, other attempts learned them in a modular manner; [Jain and Inamura \(2013\)](#) manually pre-defined local features, discretized actions for the pulling and pushing tasks, and trained a robot with a T-shaped tool. They claimed that the skills could be generalized to novel tools, though no demonstration was provided. The Keto framework ([Qin et al., 2020](#)) and the GIFT framework ([Turpin et al., 2021](#)) generated keypoints on the tools, such as grasping points and function points, based on local features. The robot then planned motion based on the keypoints. However, the keypoint approach may have difficulty on tasks where the tool contact point cannot be readily represented using only one point on the surface, such as a pencil sharpener whose contact is inside the object and the contact is more than a single point. Without using keypoints, [Abelha and Guerin \(2017\), Gajewski et al. \(2019\), Abelha et al. \(2016\)](#), and [Guerin and Ferreira \(2019\)](#) characterized the point cloud of a tool by approximating each of its segments with superquadrics and superparaboloids. They parametrized tool use with so-called p-tools, and demonstrated their technique with a wide range of tasks such as hammering and scooping in simulation or on a

physical robot. [Qin et al. \(2021\)](#) developed an integrated system and learns basic tool use, rather than pre-define the tool usages as in other studies. The system can achieve both transferable tool use and improvisatory tool use by considering both global and local geometric features. While these studies utilized visual features to transfer tool use, [Zhu et al. \(2015\)](#) included both geometric and physical features such as mass.

5.2.4 Deductive tool use

In deductive tool use, a robot should be able to utilize a novel tool to solve a task for which it has no prior knowledge. This is a very challenging task, and no current studies can perform deductive tool use to our knowledge. This type of tool use requires a robot to infer the entire affordance model, which is the relations between actions, effects, and tools, without tool use training samples as in improvisatory tool use.

5.3 Causal tool use—Multiple-manipulation tool use

Multiple-manipulation tool use involves many different types of tool use. Unlike single-manipulation tool use, the subtypes of multiple-manipulation tool use may not be interrelated. Compared with single-manipulation tool use, they may require more sophisticated manipulation skills and cognition skills such as planning, which are generally not needed in single tool use where only one tool use task is considered. In terms of tool knowledge, they usually require the full model of tool affordance knowledge.

5.3.1 Combined tool use

Combined tool use refers to using multiple tools simultaneously, such as using a fork and a knife to cut a steak. No prior studies have demonstrated combined tool use to our knowledge. We identify two main challenges of combined tool use. The first challenge is at the cognition level. A robot should choose the appropriate parameters for each tool use, such as where to cut with the knife and where to stab the steak with the fork. The second challenge is at the manipulation level, which is how to coordinate the actions of each tool. It involves collision-free motion planning and adjusting the actions of one tool based on the other tool. For example, the force exerted on the fork to stabilize the steak is dependent on the course of the cutting action with the knife. Though generating collision-free motion planning may share similar techniques in multi-agent systems [for a review, see [Rossi et al. \(2018\)](#); [Ismail et al. \(2018\)](#); [Rasheed et al. \(2022\)](#)], how to choose appropriate parameters and how to coordinate tools are issues specific to tool use and may need to be handled differently from general manipulation tasks.

5.3.2 Sequential tool use

Sequential tool use involves completing multiple tool use tasks in order. [Yamazaki et al. \(2010\)](#) designed an integrated system of daily assistive robots and applied it to the task of tidying and cleaning rooms. This system focused on failure detection and recovery, and manually defined the sequence of tasks to be completed. For a robot to be fully autonomous, the robot should be able to arrange appropriate orders and decide appropriate task parameters for each tool use task since the end state of a task is the start state of the next task.

This requirement falls under the topic of task and motion planning (TAMP) [for a review, see [Garrett et al. \(2021\)](#)]. As its name suggests, it integrates low-level motion planning which includes classic robotic manipulation techniques and high-level task planning which belongs to classic AI planning. Task planning aims to find an action skeleton to achieve a goal (e.g., pick up a pencil, use it to write, and put the pen down). Motion planning aims to find motion plans to execute in a robot (e.g., the joint states for each action). TAMP aims to find action parameters to connect task planning and motion planning (e.g., where to grasp the pencil to pick it up so the pencil can be used to write). TAMP currently has two main approaches to find action parameters: the sampling-based approach and the optimization-based approach. The sampling-based approach, which is used in the majority of TAMP studies, samples action parameters and tests the feasibility of the sampled combinations. Therefore, this approach may have difficulty when the solution space is relatively small since the probability of being able to sample the correct solution is small. In contrast, the optimization-based approach used optimization techniques such as logic-geometric programming ([Toussaint et al., 2018](#)) or sequential quadratic programming ([Hadfield-Menell et al., 2016](#)). It is able to handle problems with a small solution space more efficiently if the local optima can be handled properly. However, this approach generally requires a longer running time for tasks with many objects due to the increased dimension.

Sequential tool use has been demonstrated with optimization-based TAMP. [Toussaint et al. \(2018\)](#) enabled a robot in simulation to reach a tool that was initially out of reach with another tool in order to grab the target object. While they can handle tasks in a static environment, [Migimatsu and Bohg \(2020\)](#) improved the method with an object-centric approach to adapt to situations where objects were moved by other agents. Though this study was not demonstrated with sequential tool use, it has the potential to be applied to sequential tool use. Due to the current preliminary stage of tool use research, sequential tool use has not been demonstrated with a sampling-based approach to our knowledge.

In the above optimization-based TAMP approach, sequential tool use is only included as a demonstration to validate TAMP methods. Tool use, especially sequential tool use, usually includes multiple objects, which makes it challenging for the optimization-based approach. It is also challenging for the sampling-based approach since tool use tasks generally have a smaller solution space due to the additional constraints of tools. Therefore, alternative TAMP algorithms designed for sequential tool use may be needed due to the special requirements of tool use tasks compared with general manipulation tasks.

5.3.3 Tool selection

Tool selection is the ability to choose the most appropriate tool among many options. In order to select the most appropriate object to be used as a ram to keep a door open, [Levihn and Stilman \(2014\)](#) identified four properties of a ram. In order to learn the properties, [Wicaksono and Sammut \(2016\)](#) demonstrated a robot with an instance of the pulling task, and the robot then performed experiments to generate hypotheses about what features are important. As an example of the hypotheses, “the hook (of the pulling tool, which is a tube) must be located on the same end of tube as the cube (manipulandum).” The hypotheses are expressed in Horn clauses so that the features are qualitative. To learn the

features in a quantitative manner, Saito et al. (2018) and Brawer et al. (2020) learned the full model of tool affordances and performed tool selection.

5.3.4 Tool manufacturing

Tool manufacturing is the ability to complete a tool use task by constructing a tool by combining available materials, modifying a tool, or both. As this process may involve combining different pieces, the manipulation skills required may be similar to the skills in robotic assembly. The peg-in-hole task, which is to insert a peg in a hole, is a standard task in robotic assembly. Researchers have explored methods to improve a robot's performance, such as working with more complex parts with force-guided assembly (Dietrich et al., 2010) and increasing the speed of compliant manipulators (Bös et al., 2017) (For a review on robotic assembly with learning from demonstration, see Zhu and Hu, 2018). Beyond the peg-in-hole task, previous studies also considered the slide-in-the-groove assembly task (Peternel et al., 2015), and robot assembly that leveraged tool use such as hammering and wrenching (Gu et al., 2014).

Nair et al. (2019a,b) studied tool manufacturing by combining available parts. Their system was provided with examples of tool use, and selected appropriate parts as the grasping parts and function parts. The selection was made by comparing the similarity between the available parts and segmented parts of the demonstrated examples. The next step is to combine the parts selected with appropriate orientations. The system then performed tool use tasks to validate the assembled tool. Unlike robotic assembly, the manipulation skills required in these studies are relatively simple. It pre-defined three ways of attaching the different parts: pierce attachment, grasp attachment, and magnetic attachment. Sammut et al. (2015) designed a robot engineer to perform tool manufacturing. The robot engineer first identified important features of a tool use task, and then constructed the tool using 3D printing.

Tool manufacturing is a complicated task. The task settings of current studies reduce the difficulty of both manipulation and cognition skills. At the manipulation level, a robot may need to combine different parts with simple manipulation skills or leverage an external machine. While in animal or human tool use, the manipulation skills required in tool manufacturing are sophisticated and may even require using other tools. At the cognition level, the choice of available parts discretizes the solution space compared with the task whose solution space is continuous, such as a chimpanzee needing to make a hook to retrieve food. Tool manufacturing also requires tool affordance knowledge to identify important features of a tool to be assembled and requires planning skills to arrange the manipulation actions, especially when sequential tool use is needed.

5.4 Summary

In this section, we reviewed previous studies in robot tool use and summarized them in [Supplementary Material](#). Many studies treated tool use tasks as general manipulation tasks and focused on programming or characterizing and duplicating the actions. The learning in these non-causal tool use tasks did not consider the objects nor the effects. As a result, robots learned tool use in this manner share similar characteristics as animals performing stereotyped tool use that

the tool use demonstrate limited variations and challenge to adapt to different contexts.

As an emerging topic, relatively few studies have explored causal tool use. Early studies investigated the action-effect relations. Some studies focused on describing the effect space of an action or predicting the success or failure of an action, while other studies focused on learning how to adjust actions in order to achieve the desired effects. The latter mostly leveraged the pulling and/or the pushing tasks. To achieve transferable tool use, some studies focused on particular tasks and very few studies explored generic frameworks that may work with multiple tasks. Improvisatory tool use is even more challenging. To realize this type of tool use, some studies attempted to identify the functions of parts of a tool while others treated a tool in a holistic manner and detected key characters to represent the tool including the relations between different parts of tools. The former can generally achieve amazing results in tool use tasks whose local features are crucial to solve a task (e.g., the blade of a knife), and the latter enjoys advantages for tasks whose tools share common factors (e.g., mugs in different shapes). Though very few studies have attempted to handle both tasks. Different from these sub-types of single-manipulation tool use tasks, no previous studies have attempted deductive tool use.

Compared with single-manipulation tool use, fewer studies have addressed multiple-manipulation tool use. No previous studies reported that their systems can handle combined tool use; sequential tool use was usually treated as a regular manipulation task in TAMP; tool selection is most similar to single-manipulation tool use and gained most attention among multiple-manipulation tool use; only limited attempts have been made to address even simplified version of tool manufacturing.

6 Discussions

We defined robot tool use, provided a taxonomy of robot use, and identified the required skills in each category of tool use. As a summary, non-causal tool use focuses on the manipulation skills of using tools. Causal tool use focuses on learning or applying affordances. The sub-categories of single-manipulation tool use learn different parts of affordances. Basic tool use learns the actions-effects relation. Transferable tool use focuses on the tools-actions relation in addition to the actions-effects relation. Improvisatory tool use requires the knowledge of the full model. Deductive tool use generates affordances with general knowledge, rather than inducting the model from experiences or demonstrations. While single tool use relies on learning affordances, multiple-manipulation tool use leverages learned affordances and requires more sophisticated manipulation and/or higher-level cognition skills. In addition, we review literature on robot tool use. In this section, we discuss current or near future applications as well as future directions for robot tool use.

6.1 Current applications of robot tool use

The taxonomy can be used as a practical guideline for robot engineers when developing tool use applications. Robot engineers may start with the categorization criterion in the taxonomy since these are important features of robot tool use. This process also facilitates engineers to identify the sub-type of tool use involved in

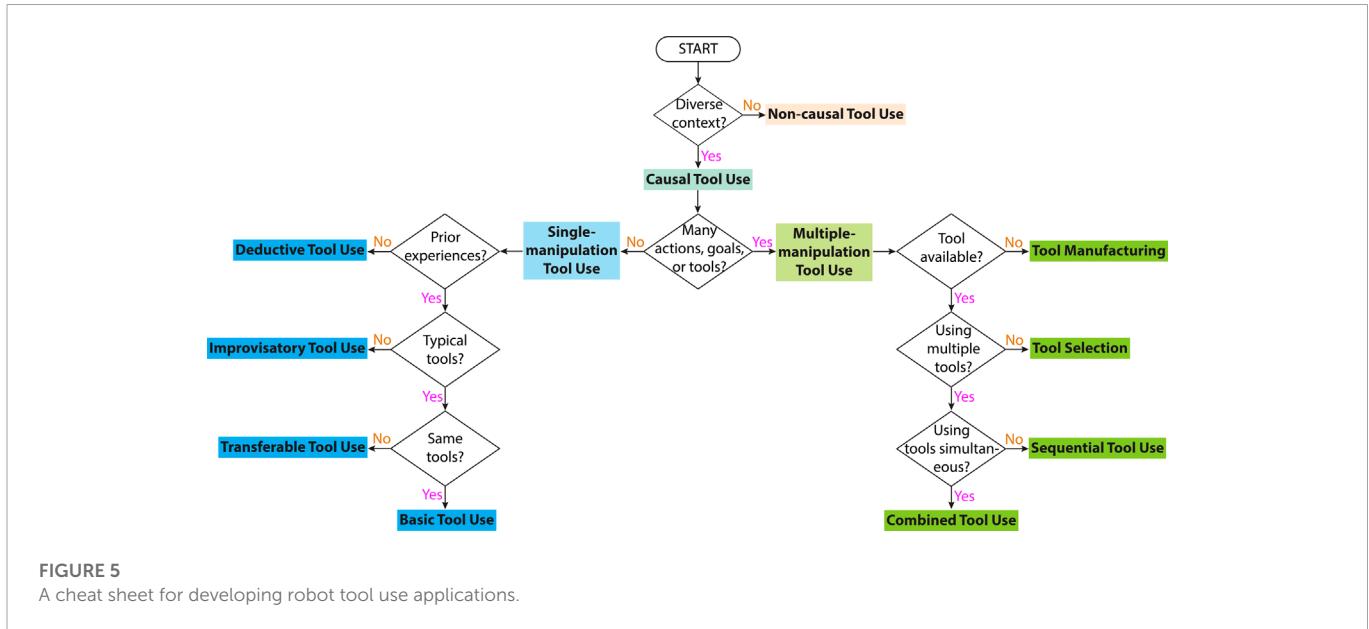


FIGURE 5
A cheat sheet for developing robot tool use applications.

the application. We include a convenient cheat sheet for engineers as **Figure 5**. Upon the identification of the sub-type, our taxonomy also provides information regarding which skills should be focused on as summarized in **Figure 4**.

We would like to emphasize that the taxonomy was not designed to decide which type of tool use is superior, but to show the differences between the sub-types of tool use. For example, causal tool use may look more appealing than non-causal tool use as the actions can be generated more flexibly in different contexts. However, the former is more challenging to learn and the performance is comprised due to the limitations of current learning techniques. For industrial settings, robots on the assembly line may be required to complete the same task repeatedly. In this scenario, being more reliable may be more important than being more flexible. Moreover, the environment is designed for the robot and it is highly-controlled and being able to adapt to diverse context is less of a concern. As a result, non-causal tool use is sufficient. Conversely, household robots may face dynamic, complex, and stochastic environments. Such environments necessitate more flexible tool use skills. Therefore, causal tool use is required. In a nutshell, what kind of tool use to implement in a certain application is dependent on the need of the task and the limitations of current techniques.

We would also like to point out that the characteristics to classify the taxonomy may not be an exhaustive list for tool use in every scenario. For example, our taxonomy is designed to consider a robot solving a tool use task by itself. However, in practice, a robot may be required to solve a task jointly with a human collaborator. Considerations around safely handling tools, and doing so in a way that fosters effective collaboration are important, but beyond the scope of the taxonomy and this survey. Moreover, different applications may have significantly different requirements making it impractical to design a guideline that may apply to every conceivable application. Our taxonomy provides a basic guideline to design a robot to complete tool use tasks, and engineers should identify specific requirements of particular applications. Not to mention that the real-world environment is noisier which adds another layer of complexity.

6.2 Open challenges of future robot tool use

The study of tool use is still in the preliminary stages, and most studies aim to solve non-causal tool use and basic tool use. We identify the following open challenges in tool use.

1. *How can a robot learn the relations between tool-manipulanda contact poses and effects in transferable tool use?* There is a lack of studies on the relationship between tool-manipulanda contact poses and tool use effects. Most studies focus on the relation between trajectories and effects.
2. *How can an integrative system for improvisatory tool use handle a wide range of tasks?* While it is challenging to improvise tool use based on either local or global features, it is even more challenging to develop a system that can solve a wide range of tool use tasks. Such a system should decide whether local or global features should be considered, or choose features beyond geometric ones.
3. *How can a robot perform deductive tool use?* The challenge for deductive tool use is the lack of prior experiences. Current techniques for other sub-types of single-manipulation tool use performs inductive reasoning that learns affordances from experiences, and cannot be applied to deductive tool use.
4. *How can a robot perform multiple-manipulation tool use?* Multiple-manipulation tool use requires a robot to perform single-manipulation tool use. In addition, each sub-type in multiple-manipulation tool use requires more sophisticated manipulation skills or higher level cognition skills. Moreover, the additional skills differ among the sub-types of multiple-manipulation tool use.
5. *How can a robot learn the dynamics in causal tool use?* It is already challenging for current studies to consider tasks that can be achieved with only kinematic control. It will be even more challenging to incorporate dynamics as it adds additional dimensions to consider.
6. *Can we design a benchmark database for standard tool use tasks?* It is not trivial to design standard tool use tasks with a benchmark database of object models to facilitate comparisons between different algorithms. The requirements of the tasks should

be detailed enough for precise replication. However, detailed requirements may lead to algorithms tailored for these tasks, and loss of generality. Moreover, it is challenging to select representative tools for improvisatory tool use as tools are expected to be used in creative manners. It is also impossible to include all possible tools for a given task due to the almost endless choices of physical objects that can be used as tools. It is also time-consuming to obtain the 3D model of an object.

7. *When and how can tool use knowledge be applied other areas in robotics?* Most studies that are relevant to tool use ignore the affordance model. For example, when learning robot grasping or robot handovers, a system typically observes how a human grasps a tool or hands over a tool, rather than inferring how a tool should be grasped or handed over based on the subsequent tasks. It is important for a system to be equipped with affordance knowledge since affordance causally determines how a tool should be grasped or handed over for to perform subsequent tool use tasks (Qin et al., 2022). However, not every study involving tool use requires a robot to learn the full affordance model, and it is important to identify which part of the model should be learned. Moreover, it also requires effort to connect tool use learning module with other modules, such as robot grasping and human-robot-collaboration tasks.

Author contributions

MQ, JB, and BS participated in writing the paper.

References

- Abelha, P., and Guerin, F. (2017). "Learning how a tool affords by simulating 3d models from the web," in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 4923–4929.
- Abelha, P., Guerin, F., and Schoeler, M. (2016). "A model-based approach to finding substitute tools in 3d vision data," in *2016 IEEE international conference on robotics and automation (ICRA)* (IEEE), 2471–2478.
- Agostini, A., Aein, M. J., Szedmak, S., Aksoy, E. E., Piater, J., and Würgütter, F. (2015). "Using structural bootstrapping for object substitution in robotic executions of human-like manipulation tasks," in *2015 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 6479–6486.
- Alcock, J. (1972). The evolution of the use of tools by feeding animals. *Evolution* 26, 464–473. doi:10.2307/2407020
- Asano, T. (1994). in *Tool using behavior and language in primates. Behavior analysis of language and cognition*. Editors S. C. Hayes, L. J. Hayes, M. Sato, and K. Ono, 145–148.
- Beck, B. B. (1980). *Animal tool behavior: The use and manufacture of tools by animals*. New York: Garland STPM Press.
- Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., and Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: Evidence from field experiments. *Anim. Cogn.* 6, 213–223. doi:10.1007/s10071-003-0183-x
- Boesch, C. (2013). "Ecology and cognition of tool use in chimpanzees," in *Tool use in animals: Cognition and ecology*, 21–47.
- Bös, J., Wahrburg, A., and Listmann, K. D. (2017). "Iteratively learned and temporally scaled force control with application to robotic assembly in unstructured environments," in *2017 IEEE international conference on robotics and automation (ICRA)* (IEEE), 3000–3007.
- Brandi, S., Kroemer, O., and Peters, J. (2014). "Generalizing pouring actions between objects using warped parameters," in *2014 IEEE-RAS international conference on humanoid robots* (IEEE), 616–621.
- Brawer, J., Qin, M., and Scassellati, B. (2020). "A causal approach to tool affordance learning," in *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 8394–8399.
- Breuer, T., Ndoundou-Hockemba, M., and Fishlock, V. (2005). First observation of tool use in wild gorillas. *PLoS Biol.* 3, e380. doi:10.1371/journal.pbio.0030380
- Byravan, A., and Fox, D. (2017). "Se3-nets: Learning rigid body motion using deep neural networks," in *2017 IEEE international conference on robotics and automation (ICRA)* (IEEE), 173–180.
- Cabrera-Álvarez, M. J., and Clayton, N. S. (2020). Neural processes underlying tool use in humans, macaques, and corvids. *Front. Psychol.* 11, 560669. doi:10.3389/fpsyg.2020.560669
- Cakmak, M., Dogar, M., Ugur, E., and Sahin, E. (2007). "Affordances as a framework for robot control," in *Proceedings of the 7th international conference on epigenetic robotics epirob'07*.
- Call, J. (2013). "Three ingredients for becoming a creative tool user," in *Tool use in animals: Cognition and ecology*, 3–20.
- Chen, H., Wan, W., and Harada, K. (2019). "Combined task and motion planning for a dual-arm robot to use a suction cup tool," in *IEEE-RAS 19th international conference on humanoid robots (humanoids)* (IEEE), 446–452.
- Chevalier-Skolnikoff, S. (1989). Spontaneous tool use and sensorimotor intelligence in cebus compared with other monkeys and apes. *Behav. brain Sci.* 12, 561–588. doi:10.1017/s0140525x00057678
- Claassens, J., and Demiris, Y. (2011). "Generalising human demonstration data by identifying affordance symmetries in object interaction trajectories," in *2011 IEEE/RSJ international conference on intelligent robots and systems (IROS)*, 1980–1985.
- Colgate, J. E., Stanley, M. C., and Brown, J. M. (1995). "Issues in the haptic display of tool use," in *Proceedings 1995 IEEE/RSJ international conference on intelligent robots and systems. Human robot interaction and cooperative robots* (IEEE), 3, 140–145.

Funding

This work is funded by the Office of Naval Research (ONR) award No. N00014-18-1-2776 and National Science Foundation (NSF) under grants No. 1955653, 1928448, 2106690, and 1813651.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frobt.2022.1009488/full#supplementary-material>

- Dehban, A., Jamone, L., Kampff, A. R., and Santos-Victor, J. (2016). "Denoising auto-encoders for learning of objects and tools affordances in continuous space," in *2016 IEEE international conference on robotics and automation (ICRA)* (IEEE), 4866–4871.
- Detry, R., Papon, J., and Mathijs, L. (2017). "Task-oriented grasping with semantic and geometric scene understanding," in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 3266–3273.
- Dietrich, F., Buchholz, D., Wobbe, F., Sowinski, F., Raatz, A., Schumacher, W., et al. (2010). "On contact models for assembly tasks: Experimental investigation beyond the peg-in-hole problem on the example of force-torque maps," in *2010 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 2313–2318.
- Dong, C., Takizawa, M., Kudoh, S., and Suehiro, T. (2019). "Precision pouring into unknown containers by service robots," in *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 5875–5882.
- Droniou, A., Ivaldi, S., and Sigaud, O. (2014). "Learning a repertoire of actions with deep neural networks," in *4th international conference on development and learning and on epigenetic robotics (IEEE)*, 229–234.
- Ek, C. H., Song, D., Huebner, K., and Kragic, D. (2010). Exploring affordances in robot grasping through latent structure representation. *Vis. Cognitive Tasks ECCV*.
- Elliott, S., and Cakmak, M. (2018). "Robotic cleaning through dirt rearrangement planning with learned transition models," in *2018 IEEE international conference on robotics and automation (ICRA)* (IEEE), 1623–1630.
- Elliott, S., Valente, M., and Cakmak, M. (2016). "Making objects graspable in confined environments through push and pull manipulation with a tool," in *2016 IEEE international conference on robotics and automation (ICRA)* (IEEE), 4851–4858.
- Elliott, S., Xu, Z., and Cakmak, M. (2017). "Learning generalizable surface cleaning actions from demonstration," in *2017 26th IEEE international symposium on robot and human interactive communication (RO-MAN)* (IEEE), 993–999.
- Fang, K., Zhu, Y., Garg, A., Kurenkov, A., Mehta, V., Fei-Fei, L., et al. (2020). Learning task-oriented grasping for tool manipulation from simulated self-supervision. *Int. J. Robotics Res.* 39, 202–216. doi:10.1177/0278364919872545
- Fitzgerald, T., Short, E., Goel, A., and Thomaz, A. (2019). Human-guided trajectory adaptation for tool transfer. *Proc. 18th Int. Conf. Aut. Agents MultiAgent Syst.*, 1350–1358.
- Forestier, S., and Oudeyer, P.-Y. (2016). "Modular active curiosity-driven discovery of tool use," in *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 3965–3972.
- Fragaszy, D., and Eshchar, Y. (2017). Tool use in nonhuman primates: Natural history, ontogenetic development and social supports for learning. *Evol. Nerv. Syst.*, 317–328. doi:10.1016/b978-0-12-804042-3.00087-7
- Gajewski, P., Ferreira, P., Bartels, G., Wang, C., Guerin, F., Indurkhy, B., et al. (2019). "Adapting everyday manipulation skills to varied scenarios," in *2019 international conference on robotics and automation (ICRA)* (IEEE), 1345–1351.
- Gao, W., and Tedrake, R. (2021). Kpam 2.0: Feedback control for category-level robotic manipulation. *IEEE Robotics Automation Lett.* 6, 2962–2969. doi:10.1109/lra.2021.3062315
- Garcia-Peraza-Herrera, L. C., Li, W., Fidon, L., Gruijthuijsen, C., Devreker, A., Attikatos, G., et al. (2017). "Toolnet: Holistically-nested real-time segmentation of robotic surgical tools," in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 5717–5722.
- Garrett, C. R., Chitnis, R., Holladay, R., Kim, B., Silver, T., Kaelbling, L. P., et al. (2021). Integrated task and motion planning. *Annu. Rev. control, robotics, Aut. Syst.* 4, 265–293. doi:10.1146/annurev-control-091420-084139
- Gemici, M. C., and Saxena, A. (2014). "Learning haptic representation for manipulating deformable food objects," in *2014 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 638–645.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton Mifflin.
- Gonçalves, A., Abrantes, J., Saponaro, G., Jamone, L., and Bernardino, A. (2014a). "Learning intermediate object affordances: Towards the development of a tool concept," in *4th international conference on development and learning and on epigenetic robotics (IEEE)*, 482–488.
- Gonçalves, A., Saponaro, G., Jamone, L., and Bernardino, A. (2014b). "Learning visual affordances of objects and tools through autonomous robot exploration," in *2014 IEEE international conference on autonomous robot systems and competitions (ICARSC)* (IEEE), 128–133.
- Goodall, J. (1964). Tool-using and aimed throwing in a community of free-living chimpanzees. *Nature* 201, 1264–1266. doi:10.1038/2011264a0
- Gu, Y., Sheng, W., and Ou, Y. (2014). "Automated assembly skill acquisition through human demonstration," in *2014 IEEE international conference on robotics and automation (ICRA)* (IEEE), 6313–6318.
- Guerin, F., and Ferreira, P. (2019). Robot manipulation in open environments: New perspectives. *IEEE Trans. cognitive Dev. Syst.* 12, 669–675. doi:10.1109/tcds.2019.2921098
- Guha, A., Yang, Y., Fermu, C., and Aloimonos, Y. (2013). "Minimalist plans for interpreting manipulation actions," in *2013 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 5908–5914.
- Hadfield-Menell, D., Lin, C., Chitnis, R., Russell, S., and Abbeel, P. (2016). "Sequential quadratic programming for task plan optimization," in *2016 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 5040–5047.
- Hoffmann, H., Chen, Z., Earl, D., Mitchell, D., Salemi, B., and Sinapov, J. (2014). Adaptive robotic tool use under variable grasps. *Robotics Aut. Syst.* 62, 833–846. doi:10.1016/j.robot.2014.02.001
- Holladay, R., Lozano-Pérez, T., and Rodriguez, A. (2019). "Force-and-motion constrained planning for tool use," in *2019 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 7409–7416.
- Hu, N., Lou, Z., Englebienne, G., and Kröse, B. J. (2014). "Learning to recognize human activities from soft labeled data," in *Robotics: Science and systems*.
- Hunt, G. R., Gray, R. D., and Taylor, A. H. (2013). "Why is tool use rare in animals," in *Tool use in animals: Cognition and ecology*, 89–118.
- Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., and Schaal, S. (2013). Dynamical movement primitives: Learning attractor models for motor behaviors. *Neural Comput.* 25, 328–373. doi:10.1162/neco_a_00393
- Ijspeert, A. J., Nakanishi, J., and Schaal, S. (2002). "Movement imitation with nonlinear dynamical systems in humanoid robots," in *Proceedings 2002 IEEE international conference on robotics and automation (cat. No. 02CH37292)* (IEEE), 2, 1398–1403.
- Ismail, Z. H., Sariff, N., and Hurtado, E. (2018). "A survey and analysis of cooperative multi-agent robot systems: Challenges and directions," in *Applications of mobile robots (IntechOpen)*, 8–14.
- Jain, R., and Inamura, T. (2013). Bayesian learning of tool affordances based on generalization of functional feature to estimate effects of unseen tools. *Artif. Life Robotics* 18, 95–103. doi:10.1007/s10015-013-0105-1
- Jamone, L., Damas, B., Santos-Victor, J., and Takanishi, A. (2013). "Online learning of humanoid robot kinematics under switching tools contexts," in *2013 IEEE international conference on robotics and automation (IEEE)*, 4811–4817.
- Jamone, L., Ugur, E., Cangelosi, A., Fadiga, L., Bernardino, A., Piater, J., et al. (2016). Affordances in psychology, neuroscience, and robotics: A survey. *IEEE Trans. Cognitive Dev. Syst.* 10, 4–25. doi:10.1109/tcds.2016.2594134
- Karayannidis, Y., Smith, C., Vina, F. E., and Kragic, D. (2014). "Online contact point estimation for uncalibrated tool use," in *2014 IEEE international conference on robotics and automation (ICRA)* (IEEE), 2488–2494.
- Katz, D., Pyuro, Y., and Brock, O. (2008). Learning to manipulate articulated objects in unstructured environments using a grounded relational representation. In *Robotics: Science and systems (citeseer)*.
- Katz, D., Venkatraman, A., Kazemi, M., Bagnell, J. A., and Stentz, A. (2014). Perceiving, learning, and exploiting object affordances for autonomous pile manipulation. *Aut. Robots* 37, 369–382. doi:10.1007/s10514-014-9407-y
- Ke, L., Wang, J., Bhattacharjee, T., Boots, B., and Srinivasa, S. (2021). "Grasping with chopsticks: Combating covariate shift in model-free imitation learning for fine manipulation," in *2021 IEEE international conference on robotics and automation (ICRA)* (IEEE), 6185–6191.
- Kemp, C. C., and Edsinger, A. (2006). "Robot manipulation of human tools: Autonomous detection and control of task relevant features," in *Proc. Of the fifth intl. Conference on development and learning*, 42.
- Kim, S.-K., Jo, J., Oh, Y., Oh, S.-R., Srinivasa, S., and Likhachev, M. (2014). "Robotic handwriting: Multi-contact manipulation based on reactional internal contact hypothesis," in *2014 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 877–884.
- Kobayashi, Y., and Hosoe, S. (2009). "Planning-space shift learning: Variable-space motion planning toward flexible extension of body schema," in *2009 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 3107–3114.
- Kober, J., Mohler, B., and Peters, J. (2008). "Learning perceptual coupling for motor primitives," in *2008 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 834–839.
- Koch, J., Büsch, L., Gomse, M., and Schüppstuhl, T. (2022). A methods-time-measurement based approach to enable action recognition for multi-variant assembly in human-robot collaboration. *Procedia CIRP* 106, 233–238. doi:10.1016/j.procir.2022.02.184
- Kokic, M., Stork, J. A., Haustein, J. A., and Kragic, D. (2017). "Affordance detection for task-specific grasping using deep learning," in *2017 IEEE-RAS 17th international conference on humanoid robotics (humanoids)* (IEEE), 91–98.
- Kormushev, P., Nenchev, D. N., Calinon, S., and Caldwell, D. G. (2011). "Upper-body kinesthetic teaching of a free-standing humanoid robot," in *2011 IEEE international conference on robotics and automation (IEEE)*, 3970–3975.
- Kroemer, O., Niekum, S., and Konidaris, G. (2021). A review of robot learning for manipulation: Challenges, representations, and algorithms. *J. Mach. Learn. Res.* 22, 1395–1476.
- Kroemer, O., Ugur, E., Oztop, E., and Peters, J. (2012). "A kernel-based approach to direct action perception," in *2012 IEEE international conference on robotics and automation (IEEE)*, 2605–2610.
- Kräutzen, M., Mann, J., Heithaus, M. R., Connor, R. C., Bejder, L., and Sherwin, W. B. (2005). Cultural transmission of tool use in bottlenose dolphins. *Proc. Natl. Acad. Sci.* 102, 8939–8943. doi:10.1073/pnas.0500232102

- Kulak, T., Silvério, J., and Calinon, S. (2020). "Fourier movement primitives: An approach for learning rhythmic robot skills from demonstrations," in *Robotics: Science and systems*.
- Kutsuzawa, K., Sakaino, S., and Tsuji, T. (2017). A control system for a tool use robot: Drawing a circle by using functions of a compass. *J. Robotics Mechatronics* 29, 395–405. doi:10.20965/jrm.2017.p0395
- Lee, D., Kunori, H., and Nakamura, Y. (2008). "Association of whole body motion from tool knowledge for humanoid robots," in *2008 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 2867–2874.
- Lee, Y.-H., and Song, K.-T. (2021). "Real-time obstacle avoidance with a virtual torque approach for a robotic tool in the end effector," in *2021 IEEE international conference on robotics and automation (ICRA)* (IEEE), 8436–8442.
- Lestel, D., and Grundmann, E. (1999). Tools, techniques and animals: The role of mediations of actions in the dynamics of social behaviours. *Soc. Sci. Inf.* 38, 367–407. doi:10.1177/053901899038003002
- Levihn, M., and Stilman, M. (2014). "Using environment objects as tools: Unconventional door opening," in *2014 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 2502–2508.
- Li, J. K., Lee, W. S., and Hsu, D. (2018). Push-net: Deep planar pushing for objects with unknown physical properties. *Robotics Sci. Syst.* 14, 1–9.
- Li, R., Pham, D. T., Huang, J., Tan, Y., Qu, M., Wang, Y., et al. (2020). Unfastening of hexagonal headed screws by a collaborative robot. *IEEE Trans. Automation Sci. Eng.* 17, 1–14. doi:10.1109/tase.2019.2958712
- Li, X., Cao, R., Feng, Y., Chen, K., Yang, B., Fu, C.-W., et al. (2022). A sim-to-real object recognition and localization framework for industrial robotic bin picking. *IEEE Robotics Automation Lett.* 7, 3961–3968. doi:10.1109/lra.2022.3149026
- Lin, Y., and Sun, Y. (2015). Robot grasp planning based on demonstrated grasp strategies. *Int. J. Robotics Res.* 34, 26–42. doi:10.1177/0278364914555544
- Lioutikov, R., Neumann, G., Maeda, G., and Peters, J. (2017). Learning movement primitive libraries through probabilistic segmentation. *Int. J. Robotics Res.* 36, 879–894. doi:10.1177/0278364917713116
- Liu, Y., Gupta, A., Abbeel, P., and Levine, S. (2018). "Imitation from observation: Learning to imitate behaviors from raw video via context translation," in *2018 IEEE international conference on robotics and automation (ICRA)* (IEEE), 1118–1125.
- Lonsdorf, E. V. (2006). What is the role of mothers in the acquisition of termite-fishing behaviors in wild chimpanzees (*pan troglodytes schweinfurthii*)? *Anim. Cogn.* 9, 36–46. doi:10.1007/s10071-005-0002-7
- Lueddecke, T., Kulpicius, T., and Woergoetter, F. (2019). Context-based affordance segmentation from 2d images for robot actions. *Robotics Aut. Syst.* 119, 92–107. doi:10.1016/j.robot.2019.05.005
- Lutscher, E., and Cheng, G. (2013). "A practical approach to generalized hierarchical task specification for indirect force controlled robots," in *2013 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 1854–1859.
- Madry, M., Song, D., Ek, C. H., and Kräig, D. (2012). "Robot bring me something to drink from": object representation for transferring task specific grasps," in *ICRA Workshop on semantic perception, mapping and exploration*, 1–6.
- Manuelli, L., Gao, W., Florence, P., and Tedrake, R. (2019). "kpam: Keypoint affordances for category-level robotic manipulation," in *The international symposium of robotics research* (Springer), 132–157.
- Mar, T., Tikhonoff, V., Metta, G., and Natale, L. (2015). "Self-supervised learning of grasp dependent tool affordances on the iCub humanoid robot," in *2015 IEEE international conference on robotics and automation (ICRA)* (IEEE), 3200–3206.
- Mar, T., Tikhonoff, V., Metta, G., and Natale, L. (2017). "Self-supervised learning of tool affordances from 3d tool representation through parallel som mapping," in *2017 IEEE international conference on robotics and automation (ICRA)* (IEEE), 894–901. doi:10.1109/ICRA.2017.7989110
- Mason, M. T. (2018). Toward robotic manipulation. *Annu. Rev. Control, Robotics, Aut. Syst.* 1, 1–28. doi:10.1146/annurev-control-060117-104848
- Matsuzawa, T. (1999). *Communication and tool use in chimpanzees: Cultural and social contexts*. Cambridge, Massachusetts: The Designs of Animal Communication, 645–671.
- Migimatsu, T., and Bohg, J. (2020). Object-centric task and motion planning in dynamic environments. *IEEE Robotics Automation Lett.* 5, 844–851. doi:10.1109/lra.2020.2965875
- Moldovan, B., Moreno, P., and van Otterlo, M. (2013). "On the use of probabilistic relational affordance models for sequential manipulation tasks in robotics," in *2013 IEEE international conference on robotics and automation (IEEE)*, 1290–1295.
- Moldovan, B., Moreno, P., Van Otterlo, M., Santos-Victor, J., and De Raedt, L. (2012). "Learning relational affordance models for robots in multi-object manipulation tasks," in *2012 IEEE international conference on robotics and automation (IEEE)*, 4373–4378.
- Montesano, L., Lopes, M., Bernardino, A., and Santos-Victor, J. (2008). Learning object affordances: From sensory-motor coordination to imitation. *IEEE Trans. Robotics* 24, 15–26. doi:10.1109/tr.2007.914848
- Montesano, L., Lopes, M., Bernardino, A., and Santos-Victor, J. (2007). "Modeling affordances using bayesian networks," in *2007 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 4102–4107.
- Muellling, K., Kober, J., and Peters, J. (2010). "Learning table tennis with a mixture of motor primitives," in *2010 10th IEEE-RAS international conference on humanoid robots (IEEE)*, 411–416.
- Murali, A., Liu, W., Marino, K., Chernova, S., and Gupta, A. (2020). "Same object, different grasps: Data and semantic knowledge for task-oriented grasping," in *Conference on robot learning*.
- Myers, A., Teo, C. L., Fermüller, C., and Aloimonos, Y. (2015). "Affordance detection of tool parts from geometric features," in *2015 IEEE international conference on robotics and automation (ICRA)* (IEEE), 1374–1381.
- Nabeshima, C., Kuniyoshi, Y., and Lungarella, M. (2007). "Towards a model for tool-body assimilation and adaptive tool-use," in *2007 IEEE 6th international conference on development and learning (IEEE)*, 288–293.
- Nabeshima, C., Lungarella, M., and Kuniyoshi, Y. (2005). "Timing-based model of body schema adaptation and its role in perception and tool use: A robot case study," in *Proceedings. The 4th international conference on development and learning, 2005 (IEEE)*, 7–12.
- Nagata, F., Watanabe, K., and Izumi, K. (2001). "Furniture polishing robot using a trajectory generator based on cutter location data," in *Proceedings 2001 ICRA. IEEE international conference on robotics and automation (cat. No. 01CH37164)* (IEEE), 1, 319–324.
- Nair, L., Balloch, J., and Chernova, S. (2019a). "Tool macgyvering: Tool construction using geometric reasoning," in *2019 international conference on robotics and automation (ICRA)* (IEEE), 5837–5843.
- Nair, L., Srikanth, N. S., Erickson, Z. M., and Chernova, S. (2019b). "Autonomous tool construction using part shape and attachment prediction," in *Robotics: Science and systems*.
- Nakamura, T., and Nagai, T. (2010). "Object concept modeling based on the relationship among appearance, usage and functions," in *2010 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 5410–5415.
- Nishide, S., Tani, J., Takahashi, T., Okuno, H. G., and Ogata, T. (2011). Tool-body assimilation of humanoid robot using a neurodynamical system. *IEEE Trans. Aut. Ment. Dev.* 4, 139–149. doi:10.1109/tamd.2011.2177660
- Oakley, K. P. (1944). Man the tool-maker. *Proc. Geologists' Assoc.* 55, 115–118. doi:10.1016/s0016-7878(44)80012-8
- Okada, K., Kojima, M., Sagawa, Y., Ichino, T., Sato, K., and Inaba, M. (2006). "Vision based behavior verification system of humanoid robot for daily environment tasks," in *2006 6th IEEE-RAS international conference on humanoid robots (IEEE)*, 7–12.
- Paraschos, A., Daniel, C., Peters, J. R., and Neumann, G. (2013). Probabilistic movement primitives. *Adv. neural Inf. Process. Syst.* 26.
- Parker, S. T., and Gibson, K. R. (1977). Object manipulation, tool use and sensorimotor intelligence as feeding adaptations in cebus monkeys and great apes. *J. Hum. Evol.* 6, 623–641. doi:10.1016/s0047-2484(77)80135-8
- Pastor, P., Hoffmann, H., Asfour, T., and Schaal, S. (2009). "Learning and generalization of motor skills by learning from demonstration," in *2009 IEEE international conference on robotics and automation (IEEE)*, 763–768.
- Pastor, P., Kalakrishnan, M., Chitta, S., Theodorou, E., and Schaal, S. (2011). "Skill learning and task outcome prediction for manipulation," in *2011 IEEE international conference on robotics and automation (IEEE)*, 3828–3834.
- Peternel, L., Petrić, T., and Babić, J. (2015). "Human-in-the-loop approach for teaching robot assembly tasks using impedance control interface," in *2015 IEEE international conference on robotics and automation (ICRA)* (IEEE), 1497–1502.
- Peters, J., and Schaal, S. (2006). "Policy gradient methods for robotics," in *2006 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 2219–2225.
- Pfeiffer, K., Escande, A., and Kheddar, A. (2017). "Nut fastening with a humanoid robot," in *2017 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE), 6142–6148.
- Qin, M., Brawer, J., and Scassellati, B. (2021). Rapidly learning generalizable and robot-agnostic tool-use skills for a wide range of tasks. *Front. Robotics AI* 8, 726463. doi:10.3389/frobt.2021.726463
- Qin, M., Brawer, J., and Scassellati, B. (2022). "Task-oriented robot-to-human handovers in collaborative tool-use tasks," in *2022 31st IEEE international conference on robot and human interactive communication (RO-MAN)* (IEEE).
- Qin, Z., Fang, K., Zhu, Y., Fei-Fei, L., and Savarese, S. (2020). "Keto: Learning keypoint representations for tool manipulation," in *2020 IEEE international conference on robotics and automation (ICRA)* (IEEE), 7278–7285.
- Raessa, M., Sánchez, D., Wan, W., Petit, D., and Harada, K. (2019). Teaching a robot to use electric tools with regrasp planning. *CAAI Trans. Intell. Technol.* 4, 54–63. doi:10.1049/trit.2018.1062
- Ramirez-Amaro, K., Beetz, M., and Cheng, G. (2014a). "Automatic segmentation and recognition of human activities from observation based on semantic reasoning," in *2014 IEEE/RSJ international conference on intelligent robots and systems (IEEE)*, 5043–5048.
- Ramirez-Amaro, K., Beetz, M., and Cheng, G. (2015). Understanding the intention of human activities through semantic perception: Observation, understanding and execution on a humanoid robot. *Adv. Robot.* 29, 345–362. doi:10.1080/01691864.2014.1003096

- Ramirez-Amaro, K., Inamura, T., Dean-León, E., Beetz, M., and Cheng, G. (2014b). "Bootstrapping humanoid robot skills by extracting semantic representations of human-like activities from virtual reality," in *2014 IEEE-RAS international conference on humanoid robots* (IEEE), 438–443.
- Rasheed, A. A. A., Abdullah, M. N., and Al-Araji, A. S. (2022). A review of multi-agent mobile robot systems applications. *Int. J. Electr. Comput. Eng.* 12.
- Robertsson, A., Olsson, T., Johansson, R., Blomdell, A., Nilsson, K., Haage, M., et al. (2006). "Implementation of industrial robot force control case study: High power stub grinding and deburring," in *2006 IEEE/RSJ international conference on intelligent robots and systems* (IEEE), 2743–2748.
- Rossi, F., Bandyopadhyay, S., Wolf, M., and Pavone, M. (2018). Review of multi-agent algorithms for collective behavior: A structural taxonomy. *IFAC-PapersOnLine* 51, 112–117. doi:10.1016/j.ifacol.2018.07.097
- Rozo, L., Jiménez, P., and Torras, C. (2013). "Force-based robot learning of pouring skills using parametric hidden markov models," in *9th international workshop on robot motion and control* (IEEE), 227–232.
- Ruiz, E., and Mayol-Cuevas, W. (2018). "Where can I do this? Geometric affordances from a single example with the interaction tensor," in *2018 IEEE international conference on robotics and automation (ICRA)* (IEEE), 2192–2199.
- Saito, N., Kim, K., Murata, S., Ogata, T., and Sugano, S. (2018). "Tool-use model considering tool selection by a robot using deep learning," in *2018 IEEE-RAS 18th international conference on humanoid robots (humanoids)* (IEEE), 270–276.
- Sammut, C., Sheh, R., Haber, A., and Wicaksono, H. (2015). "The robot engineer," in *ILP (late breaking papers)*, 101–106.
- Sarikaya, D., Corso, J. J., and Guru, K. A. (2017). Detection and localization of robotic tools in robot-assisted surgery videos using deep neural networks for region proposal and detection. *IEEE Trans. Med. Imaging* 36, 1542–1549. doi:10.1109/tmi.2017.2665671
- Schaal, S. (2006). "Dynamic movement primitives-a framework for motor control in humans and humanoid robotics," in *Adaptive motion of animals and machines* (Springer), 261–280.
- Schoeler, M., and Wörgötter, F. (2015). Bootstrapping the semantics of tools: Affordance analysis of real world objects on a per-part basis. *IEEE Trans. Cognitive Dev. Syst.* 8, 84–98. doi:10.1109/tcmd.2015.2488284
- Shao, L., Migimatsu, T., Zhang, Q., Yang, K., and Bohg, J. (2021). Concept2robot: Learning manipulation concepts from instructions and human demonstrations. *Int. J. Robotics Res.* 40, 1419–1434. doi:10.1177/02783649211046285
- Shumaker, R. W., Walkup, K. R., and Beck, B. B. (2011). *Animal tool behavior: The use and manufacture of tools by animals*. Charles Village: JHU Press.
- Sinapov, J., and Stoytchev, A. (2008). "Detecting the functional similarities between tools using a hierarchical representation of outcomes," in *7th IEEE international conference on development and learning* (IEEE), 91–96. doi:10.1109/DEVLRLN.2008.4640811
- Sinapov, J., and Stoytchev, A. (2007). "Learning and generalization of behavior-grounded tool affordances," in *2007 IEEE 6th international conference on development and learning* (IEEE), 19–24.
- Song, D., Ek, C. H., Huebner, K., and Kragic, D. (2011a). "Embodiment-specific representation of robot grasping using graphical models and latent-space discretization," in *2011 IEEE/RSJ international conference on intelligent robots and systems* (IEEE), 980–986.
- Song, D., Ek, C. H., Huebner, K., and Kragic, D. (2011b). "Multivariate discretization for bayesian network structure learning in robot grasping," in *2011 IEEE international conference on robotics and automation* (IEEE), 1944–1950.
- Song, D., Ek, C. H., Huebner, K., and Kragic, D. (2015). Task-based robot grasp planning using probabilistic inference. *IEEE Trans. robotics* 31, 546–561. doi:10.1109/tr.2015.2409912
- Song, D., Huebner, K., Kyrki, V., and Kragic, D. (2010). "Learning task constraints for robot grasping using graphical models," in *2010 IEEE/RSJ international conference on intelligent robots and systems* (IEEE), 1579–1585.
- Spagnoletti, N., Visalberghi, E., Ottoni, E., Izar, P., and Fragasy, D. (2011). Stone tool use by adult wild bearded capuchin monkeys (*Cebus libidinosus*). frequency, efficiency and tool selectivity. *J. Hum. Evol.* 61, 97–107. doi:10.1016/j.jhevol.2011.02.010
- St. Amant, R., and Horton, T. E. (2008). Revisiting the definition of animal tool use. *Anim. Behav.* 75, 1199–1208. doi:10.1016/j.anbehav.2007.09.028
- Stoytchev, A. (2005a). "Behavior-grounded representation of tool affordances," in *Proceedings of IEEE international conference on robotics and automation (ICRA)* (IEEE), 3071–3076. doi:10.1109/ROBOT.2005.1570580
- Stoytchev, A. (2003). *Computational model for an extendable robot body schema*. Tech. rep. Atlanta, Georgia: Georgia Institute of Technology.
- Stoytchev, A. (2008). "Learning the affordances of tools using a behavior-grounded approach," in *Towards affordance-based robot control* (Springer), 140–158.
- Stoytchev, A. (2007). *Robot tool behavior: A developmental approach to autonomous tool use*. Atlanta, Georgia: Georgia Institute of Technology.
- Stoytchev, A. (2005b). "Toward learning the binding affordances of objects: A behavior-grounded approach," in *Proceedings of AAAI symposium on developmental robotics* (Stanford University Menlo Park), 17–22.
- Stückler, J., and Behnke, S. (2014a). "Adaptive tool-use strategies for anthropomorphic service robots," in *2014 IEEE-RAS international conference on humanoid robots* (IEEE), 755–760.
- Stückler, J., and Behnke, S. (2014b). "Efficient deformable registration of multi-resolution surfel maps for object manipulation skill transfer," in *2014 IEEE international conference on robotics and automation (ICRA)* (IEEE), 994–1001.
- Stückler, J., and Behnke, S. (2015). "Perception of deformable objects and compliant manipulation for service robots," in *Soft robotics* (Springer), 69–80.
- Stückler, J., Droeschel, D., Gräve, K., Holz, D., Schreiber, M., Topalidou-Kyniazopoulou, A., et al. (2013). "Increasing flexibility of mobile manipulation and intuitive human-robot interaction in robocup@ home," in *Robot soccer world cup* (Springer), 135–146.
- Stückler, J., Schwarz, M., and Behnke, S. (2016). Mobile manipulation, tool use, and intuitive interaction for cognitive service robot cosero. *Front. Robotics AI* 3, 58. doi:10.3389/frobt.2016.00058
- Su, Y.-H., Huang, K., and Hannaford, B. (2018). "Real-time vision-based surgical tool segmentation with robot kinematics prior," in *International symposium on medical robotics (ISMR)* (IEEE), 1–6.
- Sukhoy, V., Georgiev, V., Wegter, T., Sweidan, R., and Stoytchev, A. (2012). Learning to slide a magnetic card through a card reader. *2012 IEEE international conference on robotics and automation (ICRA)*, 2398–2404.
- Takahashi, K., Kim, K., Ogata, T., and Sugano, S. (2017). Tool-body assimilation model considering grasping motion through deep learning. *Robotics Aut. Syst.* 91, 115–127. doi:10.1016/j.robot.2017.01.002
- Takeuchi, Y., Ge, D., and Asakawa, N. (1993). "Automated polishing process with a human-like dexterous robot," in *1993 proceedings IEEE international conference on robotics and automation (IEEE)*, 950–956.
- Tee, K. P., Cheong, S., Li, J., and Ganesh, G. (2022). A framework for tool cognition in robots without prior tool learning or observation. *Nat. Mach. Intell.* 4, 533–543. doi:10.1038/s42256-022-00500-9
- Tee, K. P., Li, J., Chen, L. T. P., Wan, K. W., and Ganesh, G. (2018). "Towards emergence of tool use in robots: Automatic tool recognition and use without prior tool learning," in *2018 IEEE international conference on robotics and automation (ICRA)* (IEEE), 6439–6446.
- Tikhonoff, V., Pattacini, U., Natale, L., and Metta, G. (2013). "Exploring affordances and tool use on the icub," in *2013 13th IEEE-RAS international conference on humanoid robots (humanoids)* (IEEE), 130–137. doi:10.1109/HUMANOIDS.2013.7029967
- Toussaint, M. A., Allen, K. R., Smith, K. A., and Tenenbaum, J. B. (2018). "Differentiable physics and stable modes for tool-use and manipulation planning," in *Robotics: Science and systems foundation*.
- Tsuji, T., Ohkuma, J., and Sakaino, S. (2015). Dynamic object manipulation considering contact condition of robot with tool. *IEEE Trans. Industrial Electron.* 63, 1972–1980. doi:10.1109/tie.2015.2508929
- Turpin, D., Wang, L., Tsogkas, S., Dickinson, S., and Garg, A. (2021). *Gift: Generalizable interaction-aware functional tool affordances without labels*. arXiv preprint arXiv:2106.14973.
- Van Lawick-Goodall, J. (1970). "Tool-using in primates and other vertebrates," in *Advances in the study of behavior* (Elsevier), 3, 195–249.
- Visalberghi, E., and Fragasy, D. (2006). What is challenging about tool use? The capuchin's perspective. *Comp. cognition Exp. Explor. animal Intell.*, 529–552.
- Vogel, J., Takemura, N., Höppner, H., van der Smagt, P., and Ganesh, G. (2017). "Hitting the sweet spot: Automatic optimization of energy transfer during tool-held hits," in *2017 IEEE international conference on robotics and automation (ICRA)* (IEEE), 1549–1556.
- Wicaksono, H., and Sammut, C. (2016). Relational tool use learning by a robot in a real and simulated world. *Proc. ACRA*.
- Wimpenny, J. H., Weir, A. A., Clayton, L., Rutz, C., and Kacelnik, A. (2009). Cognitive processes associated with sequential tool use in new caledonian crows. *PLoS One* 4, e6471. doi:10.1371/journal.pone.0006471

- Wölfel, K., and Henrich, D. (2018). "Grounding verbs for tool-dependent, sensor-based robot tasks," in *2018 27th IEEE international symposium on robot and human interactive communication (RO-MAN)* (IEEE), 378–383.
- Xie, A., Ebert, F., Levine, S., and Finn, C. (2019). "Improvisation through physical understanding: Using novel objects as tools with visual foresight," in *Proceedings of robotics: Science and systems* (Germany: FreiburgimBreisgau). doi:10.15607/RSS.2019.XV.001
- Xue, Y., and Jia, Y.-B. (2020). "Gripping a kitchen knife on the cutting board," in *2020 IEEE/RSJ international conference on intelligent robots and systems (IROS)* (IEEE). 9226–9231.
- Yamazaki, K., Ueda, R., Nozawa, S., Mori, Y., Maki, T., Hatao, N., et al. (2010). "System integration of a daily assistive robot and its application to tidying and cleaning rooms," in *2010 IEEE/RSJ international conference on intelligent robots and systems* (IEEE), 1365–1371.
- Zech, P., Haller, S., Lakani, S. R., Ridge, B., Ugur, E., and Piater, J. (2017). Computational models of affordance in robotics: A taxonomy and systematic classification. *Adapt. Behav.* 25, 235–271. doi:10.1177/1059712317726357
- Zhu, Y., Zhao, Y., and Chun Zhu, S. (2015). "Understanding tools: Task-oriented object modeling, learning and recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2855–2864.
- Zhu, Z., and Hu, H. (2018). Robot learning from demonstration in robotic assembly: A survey. *Robotics* 7, 17. doi:10.3390/robotics7020017