Inferring Higher-Order Affordances for more Natural Human-Robot Collaboration

Vasanth Sarathy
Human-Robot Interaction Laboratory, Tufts University
Medford, MA, USA
vsarathy@cs.tufts.edu

***Index Terms*—HRI; Affordances; Dempster-Shafer Theory; Uncertain Logic; Context-based Reasoning; Cognitive Robotics**

I. Introduction and Motivation

Collaborative human activities involve applying cognitive and social skills to use and manipulate objects around us, often several at a time, simultaneously and continuously. For example, cooking activities in a restaurant kitchen require cutting vegetables, monitoring the stove and keeping tools and utensils clean, all while ensuring orders are prepared and served in a coordinated and timely manner. Not only are team members recognizing various objects in their environment, but they know what to do with them (i.e., they can perceive complex object affordances). They then use these affordances to reason about the task at hand.

Being able to use objects in the environment is a highly desirable skill for robots, as well. Unfortunately, although robots are proficient at recognizing object features, they are less-skilled at recognizing what can be done with these objects.

Our research seeks to develop a computational framework for inferring affordances that can account for not only an object's physical features, but also higher-order functional, social, aesthetic, ethical and moral aspects. Towards this goal, we are developing a novel approach based on Dempster-Shafer (DS) theory [1] and first-order "uncertain logic" for inferring object affordances.

II. Background

Gibson introduced the concept of affordance to represent the relationship between an agent and its environment [2]. In cognitive science, Barsalou et al. expanded this work and attached causality to function and affordance [3]. In cognitive robotics, Montesano et al. have developed statistically-inspired causal models of affordance using Bayesian Networks (BN) to formalize the relationship between object features, actions and effects [4].

Despite these efforts, affordance inference faces many challenges that have not been overcome in the previous work. These approaches fail to provide the flexibility with which to reason about higher-order affordances (i.e., complex combinations of physical, functional and social, ethical, and aesthetic aspects) in the open world, that are influenced by changing context, social norms, historical precedence and uncertainty. For example, these current approaches cannot reason that coffee mugs afford grasping and drinking, while also affording serving as a paperweight or cupholder, or depending on the context, as family heirloom not meant to be used at all.

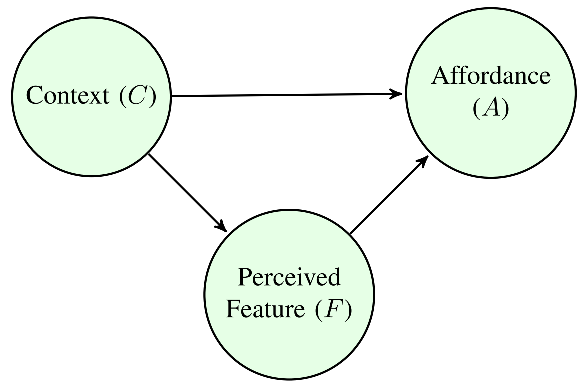

Fig. 1. Context-Sensitive Affordance Model

III. Our Computational Model

Much like the previous work, our efforts begin with Gibson's definition of affordance as a relationship between an agent and its environment. However, we diverge from the previous work in our representational and computational approach to modeling affordance [5].

We propose a model, illustrated in Fig. 1, in which an object's affordance (A) and the perceived feature of the object (F) depend on the context (C). We use Dempster-Shafer (DS) theory for inferring affordance (A) from object features (F) in contexts (C). DS theory is an uncertainty processing framework often interpreted as a generalization of the Bayesian framework. A Bayesian approach for inferring $P(A|F, C)$ by way of $P(F|A, C)$, $P(A|C)$, and $P(C)$ is not practical because we do not have a probability distribution for all the affordances for an object. Instead, we use rules taking the overall form $r :\equiv f \wedge c \implies_{[\alpha,\beta]} a$ that captures the affordance behind an object in particular contexts with $f \in F$, $c \in C$, $a \in A$, $r \in R$, $[\alpha, \beta] \subset [0, 1]$. Here, the confidence interval $[\alpha, \beta]$ is intended to capture the uncertainty associated with the affordance rule r such that if $\alpha = \beta = 1$ the rule is logically true, while $\alpha = 0$ and $\beta = 1$ assign maximum uncertainty to the rule. Rules can then be applied for a given feature percept f in given context c to obtain the implied affordance a under uncertainty about f, c, and the extent to which they imply the presence of a. We have previously shown (in the context of Indirect Speech Acts) that these types of rules are very versatile and that we can employ DS-theoretic modus ponens to make uncertain deductive and abductive inferences which cannot be made in a mere Bayesian framework [6].

IV. Human-Robot Interaction Example

Consider the example of a robotic assistant helping a human with an assembly task in which the human has asked the robot to tighten a loose screw. We would like for the robot to understand this task and the tools needed from an intuitive standpoint such that even in the absence of a screwdriver, it can reason through alternatives and find another substitute.

The robot may know of a number of rules related to its role as a helper. One rule may be: that if agent X is given a task to tighten a flat-head screw S, and X sees an object O that has a flat-head edge, then the object O has a $tightenWith$ affordance. This rule can then be represented in DS-theoretic uncertain logic a follows:

$$r^0_{[\alpha_{R_0},\beta_{R_0}]} :\equiv$$
$$hasFlatEdge(O) \wedge task(X, tighten(S, flat)) \implies$$
$$tightenWith(S, O)$$

The robot can look around the room and determine (within a certain uncertainty interval) whether or not each of the various objects that it sees has a flat edge.[1]

$$hasFlatEdge(Screwdriver)[0.95, 0.95]$$
$$hasFlatEdge(Knife)[0.9, 0.9]$$
$$hasFlatEdge(Coin)[0.75, 0.95]$$
$$hasFlatEdge(Pencil)[0, 0.95]$$

We apply DS-theoretic logical inference on rules, such as the one above, and infer uncertainties for the $tightenWith(S, O)$ affordance for each of the five objects. Based on this inference, the robot can deduce that knives and coins can be used to tighten screws in the absence of screwdrivers, but pencils cannot. Although, the rule in this example is relatively simple and primarily functional, we do contemplate scenarios that involve more complicated rules, or bundles of rules that include social and moral norms for a more complex object representation.

V. Research Approach

Our approach is to combine mathematical analysis, algorithm design, computer simulations, robotic implementation, and human-subject experiments to develop and test our computational model. In addition, we plan to study how these higher-order affordances should be learned from observation during the continuum of learning and problem-solving experiences.

As a first step, we have begun investigating a benchmarked problem of grasping objects. Past work has focused on techniques for effectively grasping objects without dropping them [8]. We are advancing this work by designing and implementing a set of context-based affordance rules to constrain the search space of possible grasp locations. For example, while a pencil can be grasped in many ways, only a few are useful for doodling, and an even fewer that conform to social norms. We plan to conduct human-subject research to then compare the predictions we can make from our knowledge base of rules against affordance perception in humans.

We also plan to investigate how robots can learn these affordance rules from observation. We plan to explore logic-based formalisms combined with a probabilistic framework to to infer commonsense rules from exploratory actions like poking and lifting. In order to represent these actions, we will study and extend existing action formalisms such as Event Calculus and Situation Calculus. In the end, we are looking to provide a much richer representation of affordance, which we believe is needed to allow robots to be adaptable to novel open-world scenarios.

VI. Conclusion

Helper robots will be critical in many sectors: helping our elderly and disabled in assisted living facilities, conducting search-and-rescue missions in unforgiving terrain to save human lives, assisting our astronauts on the space station, or even monitoring our surroundings to keep us safe from national security threats. The ultimate goal of our research is to endow robots in these critical sectors with the ability to find creative ways to use and manipulate objects, especially when there is minimal and uncertain information. We have taken the first steps towards this goal and proposed a novel approach based on Dempster-Shafer (DS) theory for inferring object affordances.

VII. Acknowledgements

This research was funded in part by ONR grant #N00014-14-1-0149.

References

[1] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

[2] J. J. Gibson, *The ecological approach to visual perception*, 1979, vol. 39.

[3] L. W. Barsalou, S. A. Sloman, and S. E. Chaigneau, "The HIPE Theory of Function," in *Representing functional features for language and space: Insights from perception, categorization and development*, 2002, vol. 30322, no. 404, pp. 1–25.

[4] L. Montesano, M. Lop, A. Bernardino, and J. Santos-Victor, "Modeling affordances using Bayesian networks," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 4102–4107. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4399511

[5] V. Sarathy and M. Scheutz, "Semantic Representation of Objects and Function," in *Proceedings of the 2015 IROS Workshop on Learning Object Affordances*, 2015.

[6] T. Williams, G. Briggs, B. Oosterveld, and M. Scheutz, "Going Beyond Literal Command-Based Instructions : Extending Robotic Natural Language Interaction Capabilities," in *AAAI*, 2015, pp. 1387–1393.

[7] L. Shapira, A. Shamir, and D. Cohen-Or, "Consistent mesh partitioning and skeletonisation using the shape diameter function," *Visual Computer*, vol. 24, no. 4, pp. 249–259, 2008.

[8] A. Ten Pas and R. Platt, "Localizing Handle-Like Grasp Affordances in 3-D Points Clouds Using Taubin Quadric Fitting," in *International Symposium on Experimental Robotics (ISER)*, 2014.

[1] For e.g., we can compute the $hasFlatEdge(O)$ predicate from 3D object meshes using Shapira's shape diameter function algorithm [7].