Towards a Prototypical Approach to Tool-Use Improvisation

Extended Abstract

Madhura Thosar
Otto-von-Guericke University Magdeburg, Germany
thosar@iks.cs.ovgu.de

Christian A. Mueller
Jacobs University, Bremen, Germany
chr.mueller@jacobs-university.de

Sebastian Zug
Technische Universitaet Bergakademie Freiberg, Germany
sebastian.zug@informatik.tu-freiberg.de

Max Pfingsthorn
OFFIS Institute for Information Technology, Oldenburg, Germany
max.pfingsthorn@offis.de

ABSTRACT

When a robot is operating in a dynamic environment, it cannot be assumed that a tool required to solve a given task will always be available. In case of a missing tool, an ideal response would be to find a substitute to complete the task. In this paper, we present a proof of concept of a grounded knowledge-based approach to tool substitution where knowledge is generated in an unsupervised manner from robot's sensory data about objects. Such robot-centric grounded knowledge is then used to identify a substitute from the available objects in the environment.

KEYWORDS

Tool Substitution; Symbol Grounding; Affordances;

ACM Reference Format:
Madhura Thosar, Christian A. Mueller, Sebastian Zug, and Max Pfingsthorn. 2019. Towards a Prototypical Approach to Tool-Use Improvisation. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 2 pages.

1 INTRODUCTION

When a robot is operating in a dynamic environment, it can not be assumed that a particular tool required to solve a task will always be available. In such scenarios, capabilities are required to mitigate the consequences of the absence of a tool by finding an alternative as humans do. This skill is significant when operating in a dynamic, uncertain environment because it allows a robot to adapt to unforeseen situations. The question is: how can a robot determine which object in the environment is a viable candidate for a substitute? It would be time consuming if a robot interacts with every single object in its environment to test its viability, which makes this approach less practical. Our approach is inspired by the way humans select a substitute for a missing tool (to a larger degree) in a non-invasive manner. For instance, when choosing a substitute between a plate and a mouse pad for a tray, humans take into account the physical and functional properties of an object [1]. A subset of these properties, however, are more relevant to enable the ability of a tray to carry objects on it such as rigidity and flatness. As a result, an ideal substitute would be the one that shares a maximum amount of relevant properties which in this case would be a plate. In this work, we propose a prototypical system named as ERSATZ (German word for substitute) where a robot-centric knowledge-driven computation is performed to identify the relevant properties of the missing tool and determine a substitute on the basis of shared relevant properties. A tool, in this work, is defined as an artifact that is designed, manufactured, and used in accordance with its designated purpose in the tasks such as hammer for hammering or boxes for storing smaller objects. Our research work primarily focuses on the selection of a substitute for a conventional tool required in the ongoing task.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

2 KNOWLEDGE ACQUISITION

The knowledge, in this work, consists of the physical and functional properties reflected in the objects. The knowledge is generated in a bottom-up manner where symbolic knowledge is generated autonomously from the perception data. This is in contrast to the top-down manner in existing knowledge bases, where the symbolic knowledge is coded by people (human-centric) and symbol grounding methods are used to ground this knowledge to a robot's perception data.

Generally, the instances of an object class tend to exhibit structural variations. Some variations will be observed in exceptional cases while some variations will be observed in majority of the cases which lead to a stereotypical understanding of an object. Therefore, the provisions should be made in conceptual knowledge about objects to model the stereotypical cases as well as exceptional cases. Moreover, the presence of the properties (physical or functional) can not be modeled simply in binary form, that is, *true* if the property is present and *false* if the property is absent. Usually, the properties are present in the object in various degrees and therefore such variations in the property measures needs to be reflected in the knowledge as well. These variations can be expressed using qualitative measures

For a bottom-up approach towards robot-centric knowledge, the primary input is measurements of a property extracted from multi-modal perception data from the individual object instances. A clustering approach is proposed to generate the symbols representing the qualitative measures. These symbols represent the degree with which a property is reflected by an object instance. Semantically, these may be interpreted for example as light, medium weight, and heavy. However, human-readable names are not given to the generated qualitative measures.

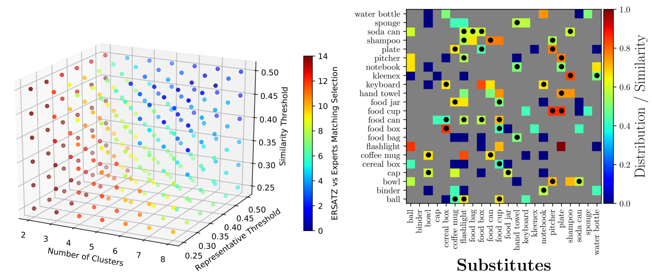
The knowledge about an object class is created by compiling the knowledge about its instances where the fuzzy set formalism is used to model the intra-class variations in the object class. A qualitative measure of a physical property is referred to as a physical quality and that of a functional property as a functional quality. The knowledge about each object instance is created by aggregating all the physical and functional qualities assigned to the object instance in the previous step. The knowledge about an object class is then generated in two steps: In the first step, the frequency of each physical/functional quality across all the instances of the object class is calculated using a *Bi-variate Joint Frequency Distribution*. In the next step, a sample proportion of each physical/functional quality in the object class is calculated. The proportion value allows to model the intra-class variations in the objects. In addition to conceptual knowledge about objects, knowledge about functional qualities, termed as *function model*, is also created. A function model consists of the frequency of each physical quality given the occurrence of a functional quality across all the object instances followed by computing a sample proportion given the frequency of the functional quality.

3 DETERMINING A SUBSTITUTE

A representative physical model and a representative functional model of an object class. consists of the physical and functional qualities, respectively, that are regarded as representative qualities of the object class if the corresponding sample proportion values are greater than a *representative model threshold*. The relevance of a representative functional quality is decided by examining whether the function model of the representative functional quality of a tool share similar physical qualities with a representative physical model of the tool. The similarity between a functional quality model and the object class of the tool is determined using the Jaccard Index. The Jaccard Index determines a similarity and dissimilarity between the two sets A and B where the similarity is calculated by dividing the magnitude of the intersection of A and B by the magnitude of the union of A and B. If the Jaccard Index is greater than the *Minimum Similarity Tolerance* threshold, then a representative functional quality is regarded as a relevant quality of the missing tool. The representative physical qualities shared by the function model of the relevant functional quality and a representative physical model of the missing tool are considered relevant to the missing tool. The substitutability of an available object is determined by computing a similarity between a representative physical model of the available object and the relevant model of the missing tool. The relevant model of the missing tool consists of its relevant physical and functional qualities.

4 EXPERIMENTS

For the experimental evaluation, we used the images from the Washington Dataset [2] to generate human-based and machine-based properties. We selected 22 object categories and for each category, random images from the given instances of the category were selected resulting into total of 692 images. The machine-centric shape-related property measurements were generated using state-of-art approach [3], which learns shape concepts from RGBD object point clouds in a data-driven and unsupervised manner. For the



(a) Impact on matching substitute selection by ERSATZ vs majority of the experts

(b) ERSATZ selections with similarity to the missing tool

physical properties *rigidity*, *weight*, *hollowness* and functional properties *support*, *blockage* and *containment*, the measurement data was generated synthetically using a human-centric intuitive model. This measurement data was used to generated knowledge about objects using k-means clustering technique. For evaluation, we generated 22 queries based on the 22 object categories, where each query consisted of a missing tool and 5 randomly selected objects as available choices for a substitute. The queries were given to 14 human experts and were asked to select a substitute in each scenario. The expert selections were aggregated and compared with the selection of ERSATZ.

The first experiment focused on tuning of the system parameters: Number of clusters, representative model threshold and minimum similarity tolerance. For the parameter *number of clusters*, the values were varied between 2 and 8. For *representative property threshold* and *acceptable similarity tolerance*, the values were varied between 0.25 and 0.50. To identify the optimum values in the experiment, we examine the effect of various permutations of the parameter values on the selection of a substitute in the 22 missing-tool scenarios and compare the result with the substitutes selected by the majority of the experts in the similar scenarios. The result is illustrated in the figure 1(a). Based on the outcomes shown in the plot, the optimal values for the parameters *number of clusters*, *representative model threshold*, and *acceptable similarity tolerance* were found to be: 4, 0.30 and 0.40 respectively. The second experiment focused on validating the substitutability of selected candidates by ERSATZ. We compared the results with expert selection and the result is illustrated using a heat plot shown in the Fig. 1(b). The grayed cells in the plots mean the corresponding object categories were not included in the available objects in the respective query. The cells that are marked with ● represents the substitutes selected by the experts and ERSATZ. Out of 22 scenarios, ERSATZ and all the experts identified the same substitutes in 20 scenarios (91%). On the other hand, the number of substitutes selected by the majority of the experts and by ERSATZ were found to be 14 (64%).

REFERENCES

[1] C. Baber. 2003. Introduction. In *Cognition and Tool Use*. Chapter 1, 1–15.
[2] Kevin Lai, Liefeng Bo, Xiaofeng Ren, and Dieter Fox. 2011. A Large-Scale Hierarchical Multi-View RGB-D Object Dataset. In *IEEE International Conference on Robotics and Automation (ICRA)*. Shanghai, China, 1817–1824.
[3] Christian A. Mueller and Andreas Birk. 2018. Conceptualization of Object Compositions Using Persistent Homology. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Madrid.