Learning Causal Relationships of Object Properties and Affordances Through Human Demonstrations and Self-Supervised Intervention for Purposeful Action in Transfer Environments

Constantin Uhde[1], Nicolas Berberich[1], Hao Ma[2], Rogelio Guadarrama[1] and Gordon Cheng[1]

***Abstract*—Learning object affordances enables robots to plan and perform purposeful actions. However, a fundamental challenge for the utilization of affordance knowledge lies in its generalization to unknown objects and environments. In this paper we present a new method for learning causal relationships between object properties and object affordances which can be transferred to other environments. Our approach, implemented on a PR2 robot, generates hypotheses of property-affordance models in a toy environment based on human demonstrations that are subsequently tested through interventional experiments. The system relies on information theory to choose experiments for maximal information gain, performs them self-supervised and uses the observed outcome to iteratively refine the set of candidate causal models. The learned causal knowledge is human-interpretable in the form of graphical models, stored in the knowledge graph. We validate our method through a task requiring affordance knowledge transfer to three different unknown environments. Our results show that extending learning from human demonstrations by causal learning through interventions led to a 71.7% decrease in model uncertainty and improved affordance classification in the transfer environments on average by 47.49%.**

***Index Terms*—Learning Categories and Concepts; Learning from Demonstration; Cognitive Modeling**

I. INTRODUCTION

KNOWLEDGE transfer from known to unknown environments through generalization is a central but challenging problem of cognitive robotics and artificial intelligence [1]. Without this ability, robotic systems need to start learning from scratch for each new environment they encounter. For humans, this challenge is literally child's play - children learn the actions they can perform with different objects through playing with toys, such as stacking blocks, carrying balls or inserting small objects into boxes. Without much visible effort, they then apply this newly acquired affordance knowledge to other scenarios such as eating breakfast. Researchers in human cognitive development have hypothesized that children achieve this generalization through *explanation* [2] [3], showing that young children between 4 and 7 years generalize significantly better when they have a mechanistic understanding of how a certain function of an object is associated with its properties. Similar to scientists [4], children identify generalizable patterns and rules ,when they play with objects (i.e. an object A can be stacked on top of an object B if A is light and small enough for grasping and both have flat surfaces). These concepts extend beyond the observation being explained [5]. Lombrozo and Carey suggested that this "Explanation for Export" functionality [6] serves the purpose of future prediction and intervention.

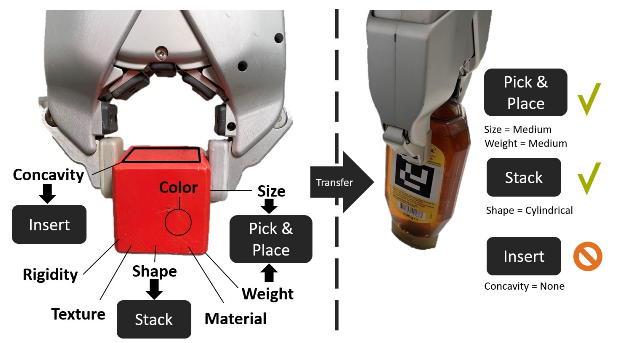


Fig. 1: Properties such as size and shape are causal for the actions a robot can perform with an object. These causal properties-affordance relationships can be learned through a combination of human demonstrations and self-supervised interventions and transferred to new environments to predict affordances of unknown objects. This causal affordance knowledge is well suited for purposeful robotic action planning.

In machine learning, which is based on generating models that represent statistical associations between features, explanations are difficult to obtain and are studied within the field of explainable artificial intelligence (XAI) [7]. While many approaches aim at making black-box machine learning models explainable, e.g. by finding training samples which have similarities with prototypical parts of the test sample [8] or attention models which provide class activation maps [9], others have suggested to work on developing inherently

Manuscript received: February, 24, 2022; Revised: May, 29, 2022; Accepted: June, 25, 2022.

This paper was recommended for publication by Editor Dana Kulic upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by Deutsche Forschungsgemeinschaft (DFG) through TUM International Graduate School of Science and Engineering (IGSSE), as well as the German Research Foundation DFG, as part of Collaborative Research Center (Sonderforschungsbereich) 1320 "EASE - Everyday Activity Science and Engineering", University of Bremen (http://www.ease-crc.org/). The research was conducted in subproject R1: NEEM-based embodied knowledge system.

[1]Constantin Uhde, Nicolas Berberich, Rogelio Guadarrama and Gordon Cheng are with Institute for Cognitive Systems, Technical University of Munich, 80333 Munich, Germany c.uhde@tum.de; n.berberich@tum.de; h.ma@tum.de; rogelio.guadarrama@tum.de; gordon.cheng@ieee.net

[2] Hao Ma is with the Technical University of Munich, 80333 Munich, Germany hao7.ma@tum.de

Digital Object Identifier (DOI): see top of this page.

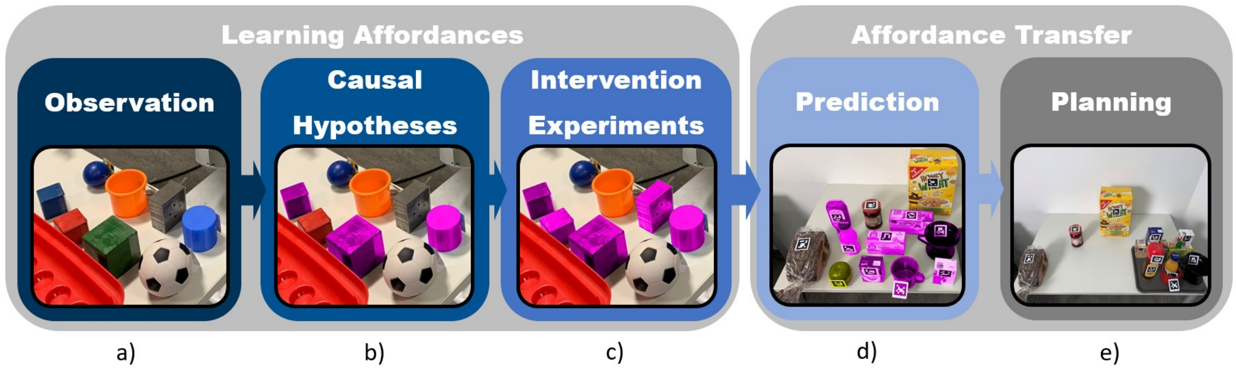


Fig. 2: Method overview illustrated on the example of learning and utilizing the affordance *"supports stacking other objects on top"*. Objects for which the affordance is classified correctly are displayed in pink, wrongly classified in yellow. From left to right, the system is confronted with a new environment and learns correlation information from human demonstrations (a). Subsequently, it creates causal hypotheses of property-affordance models based on the observations (b). The system then chooses and performs optimal interventional experiments using information theory to test and refine the model pool in a self-supervised fashion (c). This object property based affordance knowledge is transferred to a new environment with unknown object and used to predict their affordances (d). Finally, equipped with the causal affordance models, the system is able to perform a purposeful transfer planning task with previously unknown objects (e).

interpretable models instead [10].

With "causal inference", Judea Pearl and others have introduced a learning approach that is aimed at explicitly modelling the causal connections between variables which can be used for interventions and counterfactual reasoning [11], both of which are necessary for reliable robot action planning and execution.Such causal connections in the form of causal graphical models [12] can naturally serve as explanations as they represent underlying cause and effect relationships which afford to answer "why"-questions.

In order to perform purposive actions [13] towards a given goal in unstructured scenarios, robotic systems have to model their previously unknown environments and detect what actions (with their associated effects) are afforded. Especially in mission-critical settings, the robotic system should also minimize execution errors, which can be achieved by predicting the correct available actions using knowledge about causal connections in this environment. Originally discussed in psychology by James Gibson [14], knowledge about available actions can be formally represented through the concept of affordances which are relations between the object, the action and the agent's embodiment. In recent years, affordance learning has become an active field of research within robotics with varying approaches such as learning from human demonstrations, learning in simulation and learning through real-world execution [15].

In this paper we consider the following problem: A robot is situated in front of a set of toy objects and can observe human interactions with these toys, but can also interact with them itself. After a short explorative learning phase, the robot is placed into a breakfast environment with many every-day-life objects such as jam, fruit and cereal. The robot is given some time to adapt to this new scenario after which it is given goal-directed tasks such as "Clean up the breakfast table and place / stack / insert all objects on a tray."

Contribution: To solve this affordance generalization problem, we propose a new method based on learning which salient object properties such as size, weight, concavity or shape are causal requirements for a given robot action type, as depicted in Fig. 1. The system learns through the combination of observing human demonstrations, affordance hypothesis generation, hypothesis testing through deliberate experimentation based on maximizing information gain, adaptation to new environments and affordance knowledge exploitation for purposive planning. The learned causal connections are transparent and interpretable to the user, as they are stored in the searchable knowledge graph of the system.

II. RELATED WORK

The development of systems that are able to learn about object affordances is an important challenge in robotics [15][16]. Previous research has focused on predicting affordances using computer vision [17] [18]. However, good quality datasets are sparse, which some groups like Zhang et al. [19] try to address, and observational information can only used for associations in contrast to causal learning enabled by interventions [11] and neglects the central role of embodiment for robots and cognitive systems [20]. In earlier work, we demonstrated the usefulness of interventions for learning causal dependencies between actions, in order to make more profound sense of human demonstrations in a shared environment [21]. An approach on how causal inference can be applied to affordance learning has been presented in [22]. The authors focus on pushing and pulling affordances and learn structural causal models with dynamically constructed neural networks, however, similar to our previous work, they do not focus on the transfer of the learned knowledge to new environments and objects. Finally, some works addressed learning the association between object features and affordances in order to perform affordance classification like Mar et al. [23][24]. These works utilize unsupervised learning methods in order to associate object features with affordances. Both works focus on explorative discovery without learning from demonstration and don't generate a priori hypotheses about which features are salient and thus don't explicitly consider causal connections.

III. METHODS

To tackle the challenge of affordance knowledge environment transfer, we propose a robotic learning system consisting of multiple phases as illustrated in Fig. 2.

a) Observation of Human Demonstrations: Knowledge about causal dependencies can be taught and learned. Other agents may have valid models of the environment. This information is especially helpful if those agents are humans with extensive experience. In order to speed up learning causal dependency, we utilize imitation learning. Usually, this is done with video recordings of humans performing the action sequences. In this work, we created a virtual reality-based (VR) teleoperation system to let a human demonstrator interact with the objects through the robot's own embodiment. The human demonstrator is presented with the egocentric perspective of the robot, as can be seen in Fig. 3. The demonstrator's hand position is tracked in VR and corresponds to a robot end-effector goal position which is given to the robotic motion controller.

(a) Teleoperation Setup (b) Robot Observations

Fig. 3: Teleoperation setup for human demonstration for object interactions. Only a small set of demonstrations is necessary for the system to bootstrap self-supervised interventional learning.

For controlling the robotic system in an intuitive fashion, we rely on a custom implementation of a whole-body controller based on nullspace projections [25], which allows us to define hierarchical controller tasks quickly. This enables us to stay in the same embodiment for demonstrations and autonomous interventions, which simplifies the learning process by avoiding the "correspondence problem" [26].

To segment actions, we use the concept of grasp event delimitation as proposed in [27]. In order to detect grasp and release events in our system, we utilize force feedback from the PR2 grippers. After the beginning and end of an action are detected, we recognize the action type (*Carry*, *Place*, *Stack*, *Insert*) with distance thresholding, as well as spatial relations between objects. Each action has a related tool and affected object, e.g. inserting tea (tool) into a cup (affected object). Objects provide the affordance with the same name as the corresponding action, if they adhere to a certain set of properties. We selected the four affordances inspired by developmental science [16], as they seem to emerge early in development, work with primitive objects and provide a general basis for more complex affordances. We represent objects with property values, that are subjective to the embodiment of the robot (such as *light* and *heavy*). We assume that property categories are prior knowledge, inspired by [28]. Property values such as "size: large" were selected relative to the robots embodiment. The categories have been preselected in accordance to existing affordance ontology work by Bhattacharyya et al. [29]. The set of properties we consider is shown in Fig. 6. The system records the observed action sequences into a knowledge graph using the Neo4J graph data platform. We exploit the expressiveness of the declarative Neo4J querying language *Cypher* [30], in order to formulate most of the presented reasoning tasks. A reference to the conceptual structure of the graph can be seen in Fig. 4.

b) Property-Affordance Hypothesis Generation: After storing several demonstrated action sequences in the knowledge graph, we calculate correlations between property values and affordances. The intention behind this is to use the human prior in order to reduce the search space, by only looking at highly correlated values. We calculate the occurrence probability of each property value inside one property category with respect to the observed action type. Highly correlated property values get selected as candidates for causal model generation. See Eq. 1, where x_{light} is the property value light of weight X_{weight}, a_{carry} is the action carry, N_{carry}^{light} is the number of occurrences of the weight-light value in all carry actions and N_{carry} is the number of occurrences of all carry actions. The result is the occurrence probability of $X_{weight} = x_{light}$ given the affordance a_{carry}:

$$\mathbb{P}(X_{weight} = x_{light} | a_{carry}) = \frac{N_{carry}^{light}}{N_{carry}} \tag{1}$$

A high probability value means that many of the objects which were part of an observed action type share this property, making the property a likely candidate for causally affording the action. Salient candidates are selected by thresholding the probability with a user-defined value, similar to [21]. After collecting all salient property values for each the tool and the affected object components of a given affordance, we generate separate super-sets M (sets of candidate affordance models) of all possible property value combinations for both tool and affected affordance participants (e.g.

$$\begin{aligned} M_{Carry} = \{ & \{X_{weight} = x_{light,medium}, X_{size} = x_{small,medium}\}, \\ & \{X_{weight} = x_{light}, X_{size} = x_{small}\}, \\ & \{X_{weight} = x_{light}, X_{size} = x_{medium}\}, \\ & \{X_{weight} = x_{light}\}, \\ & ...\} \end{aligned}$$

for the affordance *Carry* and the affected object). We show the generated models in the set notation. Formulated in the logic notation, the model $m_{Carry} = \{X_{weight} = x_{light,medium}, X_{size} = x_{small,medium}\}$ for *Carry* would be shown as:

$$(X_{weight} = x_{light} \lor X_{weight} = x_{medium}) \land$$
$$(X_{size} = x_{small} \lor X_{size} = x_{medium}) \implies a_{carry}$$

Where a_{carry} is implied by the logical conjunction of the given property categories in an object. Note here that intra-property values are considered as logical disjunction. Each entry in the super-set is uploaded to the knowledge graph and connected to its assigned affordance with a model weight edge, as can be seen in Fig. 4. The super-set of properties-affordance models is filtered by the observed actions, so that only models remain that correctly explain the observations. These filtered models are stored in the knowledge graph and initially equally weighted, as we do not yet know which models most closely represent the causal property dependencies for the affordance.

c) Self-Supervised Experimentation through Causal Interventions: Human demonstrations give almost exclusively positive examples of which actions are possible with different

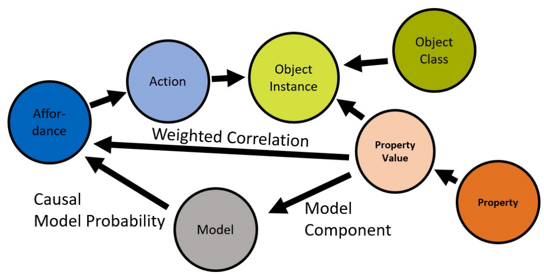


Fig. 4: Knowledge graph for learning the causal affordance models. Actions are modelled as instances of affordances and objects are based on property values. Instead of connecting object classes to affordances, our system learns causal relationships between object properties and affordances.

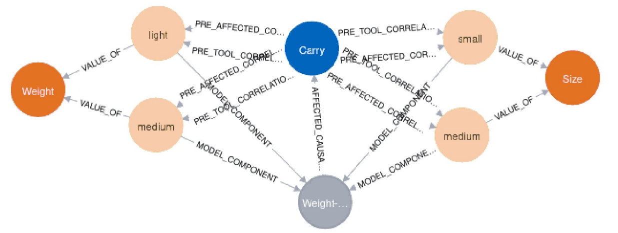


Fig. 5: A single causal model in the Neo4J graph database, depicting the ground truth for the affordance *Carry*. Each affordance has several hundred causal models initially, which get successively removed as the system learns through intervention. In a perfect experimental setup, only the ground truth model remains. If the system retains more than one model, it performs model voting to come to a classification result.

objects because humans already have an extensive concept of possible interactions and if the human fails to perform a certain action with an object, the robot can't learn from it as it doesn't know what the human intended to do. This bias towards positive examples causes our system to initially generate many candidate models. To prune these large sets of potential affordance models, the system has to discover which actions are not possible with the available objects by performing deliberate interventions. In our example model set for the affordance *Carry* presented in the last paragraph, the candidate hypothesis $m_{Carry} = \{X_{weight} = x_{light}\}$ would mean that being light is the only causal property for an object to afford to be carried. If our system performs an experiment in which it tries to carry a light, large object, but fails to do so, it can remove this hypothesis and instead assigns higher certainty weights to the hypotheses $\{X_{weight} = x_{light}, X_{size} = x_{small}\}$ and $\{X_{weight} = x_{light}, X_{size} = x_{medium}\}$. $\{X_{weight} = x_{light,medium}, X_{size} = x_{small,medium}\}$. Instead of randomly testing different interventions, we implemented a method for generating experiments with the available objects which produce a maximized information gain. We limit ourselves to generating experiments from existing objects, as the system can apply this approach iteratively, in order to learn from further exposure to new objects. This is in contrast to a setup where we generate the perfect experiment instead, by maximizing information gain over all potential property combinations, which makes the system dependent on an expert operator who provides said objects. This may also be seen as the difference between field and lab experiments [31]. We utilize the information gain formulation as seen in Eq. 2 for a possible intervention as the reduction of model uncertainty before and after the intervention. In each intervention, the robot fixes the value of one property by choosing a suitable object, e.g. fixing the weight property to *light* ($do(X_{weight} = x_{light})$), and then attempts to perform the action. Success or failure of this attempt is captured in the outcome variable y. The information gain for affordance a and intervention $do(X_{weight} = x_{light})$ is calculated as follows:

$$\begin{aligned} &Gain_a(do(X_{weight} = x_{light})) = \\ &Entropy(M_a) - \sum_{y}^{Y=\{t,f\}} \mathbb{P}(y|do(X_{weight} = x_{light}))\cdot \\ &\cdot Entropy(M_a|do(X_{weight} = x_{light})) \end{aligned} \tag{2}$$

Where $\sum_y^{Y=\{t,f\}}$ iterates through all outcomes. The entropy is calculated as $Entropy(V) = \sum_{i=1}^{nv} \mathbb{P}(v_i) \cdot \log_2(\mathbb{P}(v_i))$ where V are the current model weights M_a and the conditional model weights $M_a|do(X_{weight} = x_{light})$ respectively. The act of intervention is formalized with the *Do-Operator*, popularized by Pearl [11], which establishes the causal direction between the used random variables. In order to iteratively find the object that maximizes the information gain on the given set of candidate models, we calculate the information gain for all available objects based on their properties and test the affordance actions with the highest information gain object through an experiment. Instead of selecting among available objects, the system can also generate the overall ideal set of properties for maximum information gain. The system then requests an object with these properties for an experiment. In the remainder of this paper, we use the first approach.

After selecting an experiment (consisting of tool object, affected object and the action which is to be tested) that maximizes information gain, the system generates an action sequence using a *Planning Domain Definition Language* (PDDL) planner [32], in order to create the pre-conditions required to execute the action in question. To check whether a *red cube* affords to be stacked on another object the robot first has to grasp and carry it. For this, we first calculate possible actions based on existing affordance models and the available objects. If enough models in the knowledge graph agree on the executability of the action for the given object, it is assigned the action label, that the PDDL planner then uses to generate the setup sequence. While executing the experiment, the system uses the same action perception pipeline as for the teleoperation demonstrations to determine whether the action was successful. Thus, it employs self-supervised learning to update its affordance knowledge graph by adapting the associated causal model m_i weights for the case y_{true} based on the experiment outcomes and the Bayes rule:

$$\begin{aligned} &\mathbb{P}(m_i|do(X_{weight} = x_{light}), y_{true}) = \\ &\frac{\mathbb{P}(m_i)\mathbb{P}(y_{true}|do(X_{weight} = x_{light}), m_i)}{\mathbb{P}(y_{true}|do(X_{weight} = x_{light}))} \end{aligned} \tag{3}$$

All models whose weight converges to zero are pruned from the affordance model set M_a which is part of the knowledge graph. This process is repeated until either only one model remains or no available object combination provides sufficient information gain anymore.

d) Affordance Knowledge Transfer and Exploitation: With the refined sets of possible causal models for all affordances the system has generated abstract affordance knowledge which is independent of the specific objects in its learning

environment. This knowledge can then be transferred to new environments that include previously unknown objects. Based on the new objects' properties and the learned property-affordance models, the system predicts what actions each object affords. This decision is done by calculating a certainty value based on the vote of all related models. Given an object with the property set $o = \{Weight_{light}, Size_{medium}\}$, the model set M_{carry} from above (with the assumption that only the four shown models exist) produces a vote result of $3/4$ for the existence of the affordance *Carry* on the object o, since three out of four models agree with the object's property combination. A single such model can be seen in Fig. 5. If the certainty value lies above a user-defined threshold, the system predicts that the object affords the robot to perform the action and incorporates it into task planning. The certainty measure may be set lower or higher, depending on the required trust in the success of the planned action sequence.
The stored causal models for all affordances can be queried in a human-readable format as they adhere to the categories shown in Fig. 4. The graph database provides the declarative querying language *Cypher* for this purpose.

e) Purposeful Plan Assembly: After refining the previously generated causal affordance models, the system is able to assign probable affordances to all observed objects. Equipped with this information, a PDDL planner is able to generate an action sequence towards a given goal-condition. This plan can then be executed by the robot.

IV. RESULTS

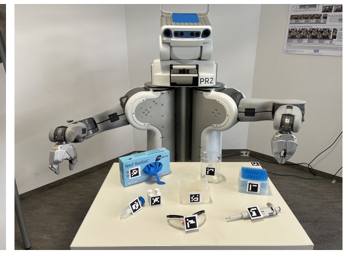
We validated the described causal affordance learning system with a knowledge transfer task between a toy objects environment and a breakfast objects environment (see Fig. 6). Similar to children, the robotic system is given the opportunity to first observe human demonstrations and then playfully learn by itself which actions are possible with different toys. Afterward, the robot is placed in front of a set of unknown breakfast objects and given the task to *place them on a tray*, requiring it to transfer and utilize its refined affordance knowledge from the toy environment. The system has to use its previously learned affordance models to infer which actions the breakfast objects afford in order to generate a goal-directed plan. The task forces the system to exploit the affordances *Stack* and *Insert* in addition to *Carry* and *Place* as there is not enough space on the tray.

The toy environment consists of 9 toy objects such as colored cubes, a cup, a cylinder and balls, with varying properties as presented in Fig. 6 a). The toys were selected to be diverse, in order to cover multiple property variations. We simplified the recognition of the objects' properties by attaching ArUco markers [33] to the objects and providing the system with a database of object-properties relations. Using the VR-teleoperation system, a human user interacted with the objects through the robot's own embodiment, while the robot observed the actions through its head cameras. The human demonstrations consisted of all affordance actions (*Carry, Place, Stack, Insert*) and covered the underlying property values, while maintaining a small set of object interactions. At the

(a) Affordance **learning** in toy environment

(b) Affordance knowledge **transfer** to breakfast environment

(c) Affordance knowledge **transfer** to bathroom environment



(d) Affordance knowledge **transfer** to chemistry environment

Environment	toy		Breakfast		Bathroom		Chemistry	
Prop. / Objects	Red Cube	Stacking Cup	Honey Wheat	Coffee Cup	Soap Dish	Hairbrush	Pipette	Beaker
Size	small	med.	large	med.	med.	med.	med.	med.
Weight	light	light	med.	light	light	med.	med.	med.
Color	red	org.	yellow	red	org.	org.	white	udef.
Material	organ.	plast.	organ.	metal	organ.	organ.	plast.	glass
Rigidity	hard	hard	med.	hard	hard	hard	hard	hard
Concavity	none	strong	none	strong	med.	none	none	strong
Surface	med.	smth.	med.	smth.	med.	med.	smth.	smth.
Shape	cub.	cyl.	cub.	cyl.	cub.	pole	pole	cyl.

(e) Subset of objects from each environment with their corresponding properties. These values are defined relative to the robot embodiment and thus represent qualitative quantities.

Fig. 6: Experimental setup for causal affordance model learning. (a) The PR2-robot is placed in front of a set of toy objects with varying properties. Through observations and interventions it learns causal connections between the object properties and actions it can perform with the objects. (b) This affordance knowledge is then transferred to a unknown breakfast environment where the robot predicts affordances for each object based on their properties. A selection of objects with their corresponding properties can be seen in table e).

beginning, the system has no information about which sets of the 52 property values are causal for each affordance. Thus, all possible combinations of them (their power set) are potential models, resulting in $4.5 \cdot 10^{15}$ candidate models (the cardinality of the power set). Instead of creating this vast amount of candidate models, the system uses the observations from the human demonstrations and only generates models which are combinations of property values that frequently occur (p>20%) in objects where the respective affordance was observed. This results in a several magnitudes smaller number of candidate causal models for each affordance as shown in Fig. 7. These model sets are then filtered with the observations such that all models are removed which do not fit to observed actions and their associated objects. E.g. if a model says that the affordance *Carry* can only be executed with an object of light weight and the observations contain a carry action with a medium weight object, the model is filtered out. For further pruning of the candidate model sets, the system generated a sequence of experiments to optimally test through interventions which

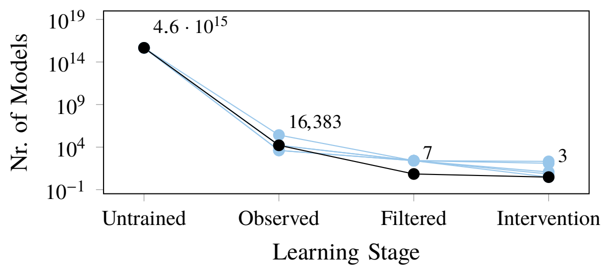


Fig. 7: Learning progress for the affordance model pools (*Carry* illustrated in black with numbers) represented in the decrease of candidate models shown on a logarithmic scale. Initially, the power set of 52 property values is considered as model candidates and subsequently pruned through observational and interventional learning.

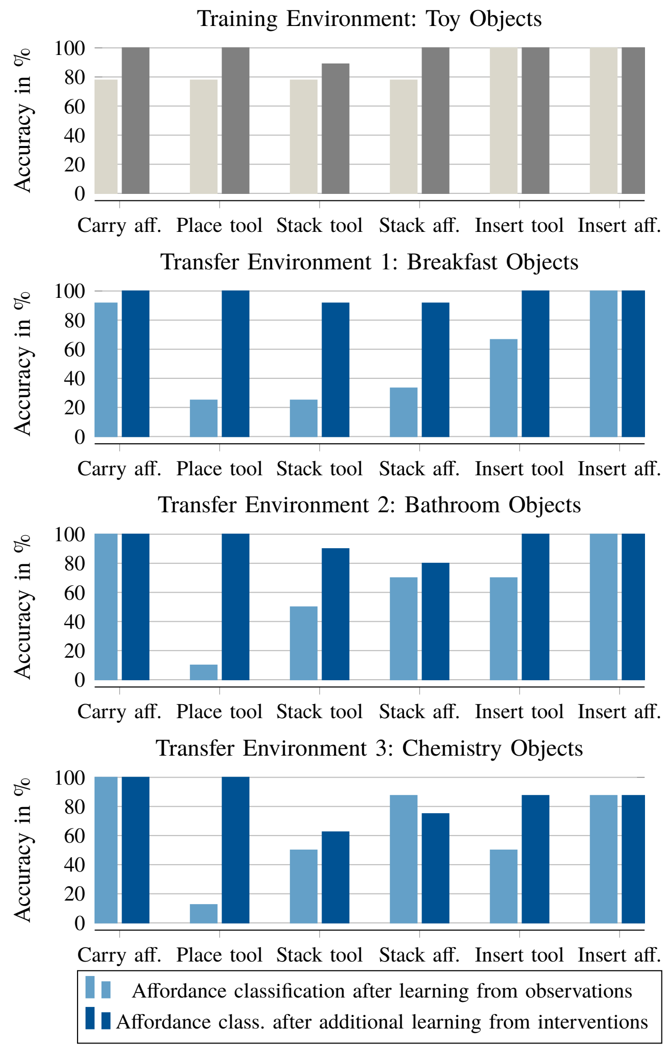


Fig. 8: Classification accuracy of object affordances with observation and intervention data from the toy object set, applied to the toy and transfer breakfast, bathroom and chemistry object sets. The bars in light blue represent the classification accuracies of the system after the observational learning phase. The dark blue bars contrast this with the classification accuracy of the intervention-trained system. A classification threshold of 0.33 has been used, meaning that 1/3 of all models in the respective model pools have to predict that the object affords the action. Affordance knowledge exclusively from observations leads to decent classification accuracy in the training environment, but generalizes badly to transfer environments. In contrast, learning causal affordance models through interventions allows the system to maintain high classification accuracy above 80% in the toy and breakfast environments, and improves the average classification accuracy by 47.49% in the three transfer environments, from 47,06% in the learning from demonstration case to 69,41% with the additional intervention stage. Ground truth data was generated from human demonstrations in the VR robot embodiment.

object properties are causal for each affordance. Instead of linearly testing all candidate models for each affordance which would have required 1282 tests, using information theory to choose experiments resulted in only 21 experiments. The PR2 robot conducted these self-supervised experiments within 25 min, requiring on average 1 min 11 s per experiment. Using the available toy objects, the system was able to reduce the candidate model sets to the following cardinalities: Carry - 3, Place - 7, Stack tool - 12, Stack affected - 3, Insert tool - 126, Insert affected - 212. To verify these remaining sets of causal models for each affordance, we applied model voting for each object to determine if the trained system correctly predicts if they afford their ground truth actions. In Fig. 8, we present accuracy values for each affordance, denoting for how many objects in both environments the system correctly classified the affordance. Finally, we test the performance of the trained system on a planning task, in which we define a carrying tray as goal for all breakfast objects the system is able to handle. The result can be seen in the last image of Fig. 2. The system correctly predicts that it can't carry the bread, cereal box and marmalade and is able to use the transferred affordance knowledge to generate a plan for fitting all other objects on the small tray by stacking them and inserting the orange juice bottle in the cup.

V. DISCUSSION

The results presented in Fig. 7 show that the system is able to effectively reduce the number of candidate affordance models through the combination of learning from observation and self-supervised interventions. This supports our initial hypothesis that interventional learning enables the system to gather information which is not available through exclusive observation, allowing it to filter out additional candidate causal models. We verify the performance in Fig. 8 by comparing the affordance classification accuracy between the purely observant and the self-supervised intervention system. The learning effect is even more pronounced when tested with unknown objects from a new environment. We argue that learning through causal interventions approximates the underlying causal model for the affordance, in contrast to observed correlations, which will only work well as long as the imitator stays and performs in the same environment. Learning about the underlying causal mechanisms lets the system transfer more effectively to novel environments. Although not the focus of this work, our system supports continual learning through performing interventional experiments in transfer environments.

In the knowledge transfer task in the breakfast environment, the robot successfully plans the placement of multiple objects on the tray (see Fig. 2 e)). It correctly assigns the *Carry* affordance and only leaves behind the objects that are too large or too . In addition to these predictions, the system provides causal reasons for them and can thus *explain* them (XAI): the box of cereal and the bread are too large and the glass of marmalade is too heavy. The system plans to stack the apple on top of the honey, even though it has no stable support. This is an effect of the residual uncertainty in the model pool, as the set of objects available for experiments is limited. To improve its classification further, the system can perform additional interventions in transfer environments. Preliminary results show that such continued model pool refinement leads to improved performance for most affordances but also highlights the remaining challenge of modelling tool-affected property interdependence, as in the case of the affordance *Insert*.

VI. CONCLUSION

In this work we demonstrated a new method for learning causal relationships between object properties and object affordances through the combination of human demonstrations and self-supervised interventions. We show that this method is able to transfer the learned affordance knowledge to an unknown scenario where it can be used for purposive action planning. The proposed method was implemented on a PR2 robot using a semantic knowledge graph and was validated on a toy environment to breakfast environment transfer task. Our results demonstrate the importance of interventions for learning domain-transferable causal property-affordance models. Generating optimal intervention experiments based on information theory led to a 71.7% decrease of candidate models and improved affordance classification in the transfer environments by 47.49%. Extending our robot-as-scientist approach, the system computes information gain to choose optimal interventions as experiments to refine its affordance knowledge. This knowledge stays always human-interpretable in the knowledge graph. In the future,we intend to address different (human) embodiments, which will allow us to compete in established affordance benchmarks. we also aim to test and improve the scalability of our system with a broader set of affordances and transfer environments, with the goal of tackling more complex real-world robotic tasks.

REFERENCES

[1] W. M. Kouw and M. Loog, "An introduction to domain adaptation and transfer learning," *arXiv preprint arXiv:1812.11806*, 2018.

[2] K. Crowley and R. S. Siegler, "Explanation and generalization in young children's strategy learning," *Child Development*, vol. 70, no. 2, p. 304–316, Mar 1999.

[3] N. Vasil, A. Ruggeri, and T. Lombrozo, "When and how children use explanations to guide generalizations," *Cognitive Development*, vol. 61, p. 101144, Jan 2022.

[4] C. Cook, N. D. Goodman, and L. E. Schulz, "Where science starts: Spontaneous experiments in preschoolers' exploratory play," *Cognition*, vol. 120, no. 3, pp. 341–349, 2011.

[5] W. V. O. Quine and J. S. Ullian, *The web of belief*. Random House New York, 1978, vol. 2.

[6] T. Lombrozo and S. Carey, "Functional explanation and the function of explanation," *Cognition*, vol. 99, no. 2, p. 167–204, Mar 2006.

[7] D. Gunning, M. Stefik, J. Choi, T. Miller, S. Stumpf, and G.-Z. Yang, "XAI—explainable artificial intelligence," *Science Robotics*, vol. 4, no. 37, p. eaay7120, Dec 2019.

[8] C. Chen, O. Li, D. Tao, A. Barnett, C. Rudin, and J. K. Su, "This looks like that: deep learning for interpretable image recognition," *Advances in neural information processing systems*, vol. 32, 2019.

[9] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *IEEE conference on computer vision and pattern recognition*, 2016, pp. 2921–2929.

[10] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, p. 206–215, May 2019.

[11] J. Pearl, M. Glymour, and N. P. Jewell, *Causal inference in statistics: A primer*. John Wiley & Sons, 2016.

[12] C. Glymour, K. Zhang, and P. Spirtes, "Review of causal discovery methods based on graphical models," *Frontiers in genetics*, vol. 10, p. 524, 2019.

[13] G. Cheng, K. Ramirez-Amaro, M. Beetz, and Y. Kuniyoshi, "Purposive learning: Robot reasoning about the meanings of human activities," *Science Robotics*, vol. 4, no. 26, p. eaav1530, 2019.

[14] J. J. Gibson, "The theory of affordances," *Hilldale, USA*, vol. 1, no. 2, pp. 67–82, 1977.

[15] N. Yamanobe, W. Wan, I. G. Ramirez-Alpizar, D. Petit, T. Tsuji, S. Akizuki, M. Hashimoto, K. Nagata, and K. Harada, "A brief review of affordance in robotic manipulation research," *Journal of the Robotics Society of Japan*, vol. 36, pp. 327–337, 01 2018.

[16] "Affordances in psychology, neuroscience, and robotics: A survey," vol. 10, p. 4–25, Mar 2018.

[17] T. Lüddecke and F. Wörgötter, "Learning to label affordances from simulated and real data," p. 11.

[18] D. Katz, A. Venkatraman, M. Kazemi, J. A. Bagnell, and A. Stentz, "Perceiving, learning, and exploiting object affordances for autonomous pile manipulation," p. 8.

[19] L. Zhang, W. Du, S. Zhou, J. Wang, and J. Shi, "Inpaint2learn: A self-supervised framework for affordance learning," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2022, pp. 2665–2674.

[20] D. Vernon, G. Metta, and G. Sandini, "Embodiment in cognitive systems: on the mutual dependence of cognition & robotics," *cognition*, vol. 8, no. 11, p. 17, 2010.

[21] C. Uhde, N. Berberich, K. Ramirez-Amaro, and G. Cheng, "The robot as scientist: Using mental simulation to test causal hypotheses extracted from human activities in virtual reality," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 8081–8086.

[22] J. Brawer, M. Qin, and B. Scassellati, "A causal approach to tool affordance learning," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 8394–8399.

[23] T. Mar, V. Tikhanoff, and L. Natale, "What can i do with this tool? self-supervised learning of tool affordances from their 3-d geometry," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 3, pp. 595–610, 2017.

[24] T. Mar, V. Tikhanoff, G. Metta, and L. Natale, "Self-supervised learning of grasp dependent tool affordances on the icub humanoid robot," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 3200–3206.

[25] L. Sentis and O. Khatib, "A whole-body control framework for humanoids operating in human environments," in *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.*, 2006, pp. 2641–2648.

[26] C. L. Nehaniv, K. Dautenhahn, *et al.*, "The correspondence problem," *Imitation in animals and artifacts*, vol. 41, 2002.

[27] K. Pastra and Y. Aloimonos, "The minimalist grammar of action," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 367, no. 1585, pp. 103–117, 2012.

[28] G. Vallortigara, "Core knowledge of object, number, and geometry: A comparative and neural approach," *Cognitive neuropsychology*, vol. 29, no. 1-2, pp. 213–236, 2012.

[29] R. Bhattacharyya, Z. Bhuyan, and S. M. Hazarika, "O-pro: an ontology for object affordance reasoning," in *International Conference on Intelligent Human Computer Interaction*. Springer, 2016, pp. 39–50.

[30] N. Francis, A. Green, P. Guagliardo, L. Libkin, T. Lindaaker, V. Marsault, S. Plantikow, M. Rydberg, P. Selmer, and A. Taylor, "Cypher: An evolving query language for property graphs," in *Proceedings of the 2018 International Conference on Management of Data*, 2018, pp. 1433–1445.

[31] J. M. Diamond, "Ecology: Laboratory, field and natural experiments," *Nature*, vol. 304, no. 5927, pp. 586–587, 1983.

[32] Y. Alkhazraji, M. Frorath, M. Grützner, M. Helmert, T. Liebetraut, R. Mattmüller, M. Ortlieb, J. Seipp, T. Springenberg, P. Stahl, and J. Wülfing, "Pyperplan," https://doi.org/10.5281/zenodo.3700819, 2020. [Online]. Available: https://doi.org/10.5281/zenodo.3700819

[33] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, 2014.