Active Learning for Teaching a Robot Grounded Relational Symbols

Johannes Kulick and **Marc Toussaint** and **Tobias Lang** and **Manuel Lopes**

Universität Stuttgart
Universitätsstr. 38
70569 Stuttgart
johannes.kulick@ipvs.uni-stuttgart.de
marc.toussaint@ipvs.uni-stuttgart.de

Freie Universität Berlin
Arnimallee 7
14195 Berlin
tobias.lang@fu-berlin.de

FLOWERS Team
INRIA Bordeaux Sud-Ouest, Bat. A29
351 Cours de la Liberation
33405 Talence
manuel.lopes@inria.fr

Abstract

We investigate an interactive teaching scenario, where a human teaches a robot symbols which abstract the geometric properties of objects. There are multiple motivations for this scenario: First, state-of-the-art methods for relational reinforcement learning demonstrate that we can learn and employ strongly generalizing abstract models with great success for goal-directed object manipulation. However, these methods rely on *given* grounded action and state symbols and raise the classical question: *Where do the symbols come from?* Second, existing research on learning from human-robot interaction has focused mostly on the motion level (e.g., imitation learning). However, if the goal of teaching is to enable the robot to autonomously solve sequential manipulation tasks in a goal-directed manner, the human should have the possibility to teach the relevant abstractions to describe the task and let the robot eventually leverage powerful relational RL methods. In this paper we formalize human-robot teaching of grounded symbols as an active learning problem, where the robot actively generates pick&place geometric situations that maximize its information gain about the symbol to be learned. We demonstrate that the learned symbols can be used by a robot in a relational RL framework to learn probabilistic relational rules and use them to solve object manipulation tasks in a goal-directed manner.

1 Introduction

Complex object manipulation tasks require both motion generation on the geometric level as well as sequential composition and reasoning on more abstract, e.g. symbolic relational representations [Lemaignan *et al.*, 2011; Katz *et al.*, 2008]. In existing systems that incorporate both levels it is usually assumed that a set of grounded action symbols (motion/manipulation primitives) as well as state symbols are predefined [Beetz *et al.*, 2010]. In manipulation scenarios where the world is composed of objects and the state is naturally described by properties and relations of objects, relational representations based on conjunctions of logical predicates are well suited [Džeroski *et al.*, 2001]. This led to the recent development of efficient methods for relational reinforcement learning (RL) and planning, which combine probabilistic learning and planning with relational representations [Lang and Toussaint, 2010]. However, the learning of appropriate action and state symbols themselves remains a fundamental challenge.

Instead of predefining a set of symbols by the system designer it is desirable to teach a robot new symbols, that enable him to reason about new situations, during its lifetime. Instead of aiming for a full autonomous learning of symbols we consider learning from human-robot interaction. A teacher labels a few examples of situations that are descriptive for a new symbol and considered to be useful for a new task. From this the robot learns a first model of the symbol which it actively aims to refine.

Concretely, we propose to frame the problem as an active learning problem [Settles, 2009], which generally means that the learning system incrementally chooses data points that promise the maximal learning effect, e.g. reduction in uncertainty, for its current model. In our case this implies that the robot actively generates feasible geometric situations by moving objects in a way that maximizes the expected information gain of its current symbol model. Each such query is answered by the human on whether the symbol in question is true or false, leading to an update of the robot's symbol model. This mimics the behavior of infants or children, which do not come up with symbols completely on their own, but use the help of teachers, like parents or siblings by asking them questions.

The active learning strategy is to ensure that only few symbol demonstrations are needed – the human should not have the impression that the robot is "asking" (by physically generating geometric configurations) redundant questions about the symbol (see Fig. 1). To enable active learning we formalize symbol models as probabilistic classifiers on non-linear features. These also capture the uncertainty of the model and allow for choosing geometric queries that promise high information gain.

In the next section we will review related work. Sec. 3 introduces our formal model to ground relational symbols in geometric features of the physical world and shows how a robot can actively learn such grounded symbols. Sec. 4 presents our empirical evaluation both in simulation and on a real robot

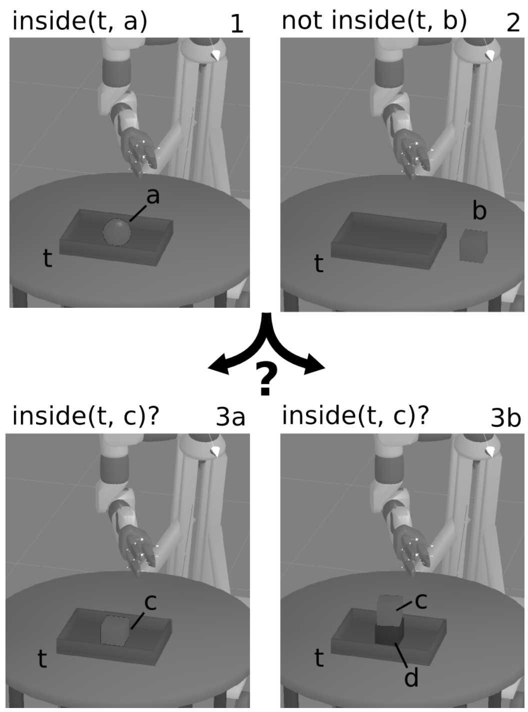


Figure 1: In active learning of grounded relational symbols, the robot generates situations in which it is uncertain about the symbol grounding. After having seen the examples in (1) and (2), the robot can decide whether it wants to see (3a) or (3b). An actively learning robot takes its current knowledge into account and prefers to see the more novel (3b).

which demonstrates the efficiency of active symbol learning.

2 Related Work

The symbol grounding problem is well discussed in artificial intelligence. Both Searle with the Chinese room problem [Searle, 1980] and Harnad [Harnad, 1990] worked on this question rather philosophically and under strict theoretical boundaries, especially the demand of no semantic commitment. This practically forbids to give a system any information about the symbols and hindered progress for many years [Taddeo and Floridi, 2005].

An alternative approach is to aim for symbols that allow systems to communicate information to other agents, as in Steels' language games [Steels, 2001]. In contrast to these approaches, we aim at symbols that are acquired from and support human-robot interaction and can eventually be used in a reinforcement learning setup.

Full-fledged language acquisition is a complex process of cross-situational learning [Siskind, 1996; Frank *et al.*, 2008]. Existing work has proposed methods to learn to associate words to specific contexts by observing cross-situational robot-object interaction descriptions [Salvi *et al.*, 2012] or by learning how to interpret natural language navigation from observations [Chen and Mooney, 2011]. We do not address the general problem of language acquisition and instead focus on the explicit human-robot teaching of symbols.

In robotics, symbol anchoring is typically hand-coded; examples for such approaches include [Shapiro and Ismail, 2001; Beetz *et al.*, 2010; Katz *et al.*, 2008]. Our approach aims to avoid the necessity to hand-code symbols for robotic planning and symbolic reinforcement learning by learning such symbolic descriptions of contexts from user input.

Active learning is a technique that actively chooses training data points which are optimal under certain information seeking criteria, to minimize the number of queries needed. Several criteria have been used to create queries in an optimal manner. An overview over different active learning approaches can be found from [Settles, 2009]. A statistically active learning technique was proposed by [Cohn *et al.*, 1996]. Our approach will use active learning to reduce the amount of training data and thereby limit the humans effort in teaching the robot symbols. Unlike existing active learning scenarios, *actively choosing* a training data point here requires the robot to literally pick-and-place objects to generate an informative geometric situation and query the human for a symbol label.

Active learning in a robotic context recently raised interest of researchers [Cakmak and Thomaz, 2012; Cakmak *et al.*, 2010; Chao *et al.*, 2010]. The focus of this research is to enable technically not trained people to teach robots new skills. Active learning techniques are therefore used to ask questions considered *good*.

In summary, our contributions over existing work can be characterized as follows: (1) the formalization of symbol models in terms of probabilistic classifiers that allow for the estimation of symbol uncertainty; (2) active learning applied to the problem of symbol learning from human-robot interaction; (3) methods to physically generate informative and feasible queries by robot object manipulation.

3 Active Symbol Learning from Human Interaction/Teaching

We will describe our methods for active symbol learning from interaction with a human on several levels. First we define how we model symbols using Gaussian Process Classification and how symbols can be trained given labeled data. Second we move to the active learning case where the robot actively queries the human to provide labeled data. Here a major issue arises due to the specific robotic context: Actively generating queries requires the robot to actively generate physically feasible geometric constellations of objects which are at the same time informative in the active learning sense.

3.1 Modeling and Learning Symbols

We assume that a scene s is composed of objects. For each object i we have a geometric feature vector $x_i \in \mathbb{R}^m$ describing geometric properties of the object (like radius, height, absolute position, color). Likewise, for each pair of objects (ij) we have features $x_{ij} \in \mathbb{R}^M$ describing the objects' geometric relation (like the difference $x_i - x_j$, distance, relative position, etc.).

We define a grounded relational symbol $\sigma = (p, f)$ as a tuple of a first order predicate p and a discriminative function f grounding the symbol. f determines the probability that the predicate p holds for objects $\mathbf{o} = (o_i, o_j)$ given their features

$x_{\mathbf{o}} = (x_i, x_j, x_{i,j})^T$ (or $\mathbf{o} = o_i$ and $x_{\mathbf{o}} = x_i$ for the unary case) in state s,

$$P(p(\mathbf{o}) \mid s) = \text{sig}(f(x_{\mathbf{o}})) \ . \tag{1}$$

where sig is the logistic function $\text{sig}(z) = \frac{1}{e^{-z}+1}$. For each continuous state s we can now define a symbolic state $t \in \{0,1\}^v$. For all objects and combinations of objects $\mathbf{o}$ and symbols $\sigma = (p, f)$, t has an entry t_i which is 1 if and only if $\text{sig}(f(x_{\mathbf{o}})) > 0.5$ and 0 otherwise. t describes which symbolic predicates are considered to be true in state s.

We propose using a Gaussian Process Classification (GPC) to learn f. GPC uses a gaussian process to represent the predictive distribution $P(f(x) \mid D)$ over the discriminative function $f(x)$, with D being the observed data [Rasmussen and Williams, 2006] (Section 3.3). GPC can also be understood as Bayesian Kernel logistic regression, where the class probability $P(y(x)|D) = \int_{f(x)} \text{sig}(f(x)) \ P(f(x)|D) \ df$ is given by the logistic sigmoid of the discriminative function, but now taking the expectation w.r.t. the posterior $P(f(x)|D)$. The MAP discriminative function $f^*(x)$ of the Gaussian process is the same as the one for standard Kernel logistic regression. Since the gaussian process models $P(f(x)|D)$ as Gaussian, the computation of the exact class probabilities $P(y(x) \mid D)$ requires to evaluate the convolution of the logistic function with a Gaussian—which is standardly approximated using the probit function $\varphi(x) = \int_{-\infty}^{x} \mathcal{N}(0,1) \ dx$.

3.2 Active learning by generating physically feasible informative situations

Estimating the reduction in predictive uncertainty. The robot should generate situations x such that the new sample (x, y) minimizes the future generalization error of the estimated grounding — in our classification case the uncertainty of the class labels $P(y(x)|D)$. We first address the estimation of this reduction in predictive uncertainty.

We define $\upsilon_D(x) = H(y(x)|D)$ as the entropy of the predictive class distribution $P(y(x)|D)$; note that entropy corresponds to expected neg-log-likelihood, i.e., expected "error". We define the overall predictive uncertainty Υ of our Bayesian classifier $P(y(x)|D)$ as integrated entropy

$$\Upsilon(D) := \int \upsilon_D(x) \ \mathrm{d}x \ . \tag{2}$$

For active learning, the system chooses the sample $\tilde{x}$ with

$$\tilde{x} = \underset{x'}{\operatorname{argmax}} \quad \Upsilon(D) - E_{y'} \left[\Upsilon(D \cup \{(x', y')\}) \right] , \tag{3}$$

where y' is the label returned by the teacher for query x' and $E_{y'}[\cdot]$ is the expectation over y'. Since y' is unknown before we actually ask the teacher, we have to estimate the change in predictive uncertainty with the expectation over all possible labels.

The predictive distribution $P(y(x)|D \cup \{(\tilde{x}, \tilde{y})\})$ learned from $D \cup \{(\tilde{x}, \tilde{y})\}$ including the chosen sample $\tilde{x}$ has the lowest uncertainty in comparison to $P(y(x)|D)$ learned from D only. Typically, the integral in equation (2) as well as the argmax for $\tilde{x}$ cannot be computed analytically. To approximate the integral (2) we perform a Monte-Carlo integration, by sampling k physically feasible reference configurations. This is a standard approach to estimate the global uncertainty [Cohn *et al.*, 1996]. The argmax in (3) is approximated using the pooling approach explained in the next paragraph.

While this approximation leads to good results it is computational expensive, since υ_D has to be computed at all reference points for all tested samples. This is especially problematic in situations where time for an oracle is considered expensive, such as in experiments with a human teacher.

Therefore we compare this objective with a local estimation of the uncertainty. For this purpose we use the entropy of the predictive distribution directly as optimization criterion and the robot chooses the sample $\hat{x}$ with

$$\hat{x} = \underset{x'}{\operatorname{argmax}} \quad \upsilon_D(x') \ . \tag{4}$$

This criterion scales well to high dimensional spaces. The optimization is also done via a pooling approach.

Sampling physically feasible & informative situations

Given a classifier f we can estimate the reduction of predictive uncertainty for any new input y. However, the inputs y to a classifier need to be features of a real physical situation: Not every feature vector can be generated by a physically feasible situation and the feature map $\phi : x \rightarrow y$ cannot be inverted to retrieve a physical situation that generates a given feature. The physical feasibility can be viewed as a structural constraint of the classifier's input domain which is not present in typical active learning approaches and which we need to explicitly take into account. Thus we need to only sample from this subspace during learning the groundings. To cover this subspace we use a physical simulator to generate a large set of physically feasible situations, which are steady. The robot now can generate the features of the situations from the simulator and compute the expected reduction of the predictive uncertainty from this simulated situation.

This approach is a form of pool based active learning. Pooling assumes that large sets of unlabeled data can be retrieved cheaply. By using the simulator we do not need to actually build the situations in real world but can generate them fast and still have a notion of steadiness and feasibility. However, the actual sample is then evaluated in real-world.

Robot manipulation to generate informative situations

To actually interact with a human teacher in the given active learning scenario the agent needs to manipulate objects in the real world and literally generate informative situations. Because we compute the features from simulated situations we have access to the positions of the objects and build the most interesting scene in real world. But since the sensorimotor loop of a robot is subject of many sources of noise, we are not able to precisely generate the same situation as we had in the simulation. While the normal pool based active learning assumes that the samples can directly be labeled by the teacher, the robot-human scenario introduces another step in the processing and the labeled sample y' is different from the originally generated sample y. It does not necessarily lead to the same information gain.

To place objects into desired poses we generate robot pick-and-place motions using standard robot trajectory optimization methods. Concerning the object picking, the pregrasp (hand and finger poses before closing fingers) and reaching trajectory to the pregrasp are jointly optimized; the objective function includes cost terms for trajectory length and collision and joint limit avoidance during the trajectory as well as cost terms for the relative wrist-to-object pose, finger tips to surface distances, and opposedness of finger tips of the final pregrasp. The optimization is done with a fast sequential 2nd order optimizer (Levenberg-Marquardt-damped, iterated Gauss-Newton) that exploits the sequential structure of the trajectory in the (banded) matrix inversion. The placing trajectories are similarly optimized, with cost terms reflecting the desired final placement of the object.

4 Evaluation

We investigated whether (a) our formal model of symbol grounding allows to ground relational symbols in the physical world, (b) our active learning method permits faster learning than random learning, and (c) the learned symbols enable abstract robot learning and reasoning, opening the door to the techniques developed in symbolic artificial intelligence.

4.1 Experiment setup

We examine our symbol learning approach in a robot manipulation scenario where a robot manipulates balls, cubes and cylinders on a table. The robot can execute three motor primitives: *closeHandAround(X)* grabs an object; *openHandOver(X)* opens the robot's hand above some other object and *openHandAt(X)* opens the hand at a specific point in space. While opening the hand always succeeds, grabbing objects might fail (both in simulation and real world).

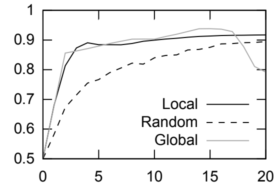
The robot can sense objects through a depth-sensor perception system, which performs the task of objects segmentation and recognition with a point-cloud based software system. For the task of this paper the perception system is considered to be a black box. The object observations are described by continuous feature vectors $x_i \in \mathbb{R}^4$ comprising the 3-dimensional position of the center and the size of object o_i. The object features are used to construct features $x_{ij} = x_i - x_j$ describing the relationship between objects, namely the distance and size differences and the sine of the angle between the main axes of the objects.

We performed experiments both in simulation (see Fig. 4) and on a real robot (see Fig. 3). In both scenarios the robot consists of a Schunk Light Weight arm with 7 DoF, a Schunk Dexterous Hand with 7 DoF, 6x14 tactile arrays on each of the 6 finger segments and a Microsoft Kinect depth sensor.

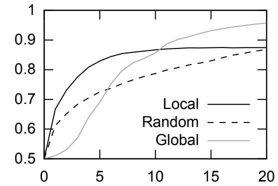
4.2 Experiment 1: Quantitative results in simulation

In the first experiment we let an agent learn unary and binary spatial relations of objects in a simulated environment. For each symbol the agent is provided with an example where the relation underlying the symbol holds.

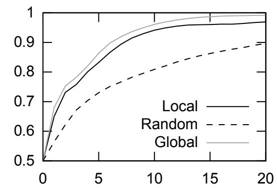
The unary predicate, the robot should learn, is *upright(X)* while the binary predicates in this experiment are *on(X,Y)* and *close(X,Y)*.



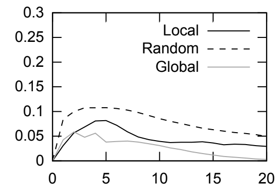
(a) $upright(\cdot)$



(b) $close(\cdot, \cdot)$



(c) $on(\cdot, \cdot)$



(d) Standard deviation of *on* learner

Figure 2: *Experiment 1: Comparison of the proposed active learning approach with random learning.* The results for learning unary and binary grounded relational symbols in simulation show that active learning outperforms random learning. (d) shows the deviation of the learner *not* the deviation of the mean estimator. This deviation is very small due to the high number of experiments ($n = 1000$).

The agent should now learn a predictive model for the given symbol by querying a teacher with new situations. The teacher in these experiments is a handcrafted oracle, that computes whether a relation holds or not. To acquire quantitative data we do these experiments without a human in the loop. Experiments in the human-robot interaction research have shown that asking good questions is important, when interacting with human teachers [Cakmak *et al.*, 2010], especially if not technically trained, to keep the attention of the teacher. So one might expect slightly different results with a human teacher. The results of our first experiment should be considered to be a baseline for further investigation.

The agent actively queries 20 samples from the teacher. After each query the agent can recompute its model and then choose the next query-point. We call this a run. We performed 1000 runs each for both active learning criteria (i.e. the local and the global, see Sec. 3.2) and a random learner (i.e. a learner choosing random samples).

To evaluate the learning rate the classifier is tested with 5000 random samples after each query and the classification rate is computed. Note that this is simply for evaluation of the classifier. We do not provide the learner with 5000 labeled examples. In Fig. 2 we show the results of these tests. In Fig. 2(a) - 2(c) we show the mean classification rate of the 1000 experiments after each query and Fig. 2(d) shows the standard deviation of one learner. The standard deviation of the mean estimator is very small (< 0.01 for all points), due to the high number of experiments. Hence we do not show it in the graphs.

It can be seen that in all cases the active learning approach outperforms the random learning. Also the standard deviation is smaller. When comparing the different information gain measurements, one can see that the learning progress is lower

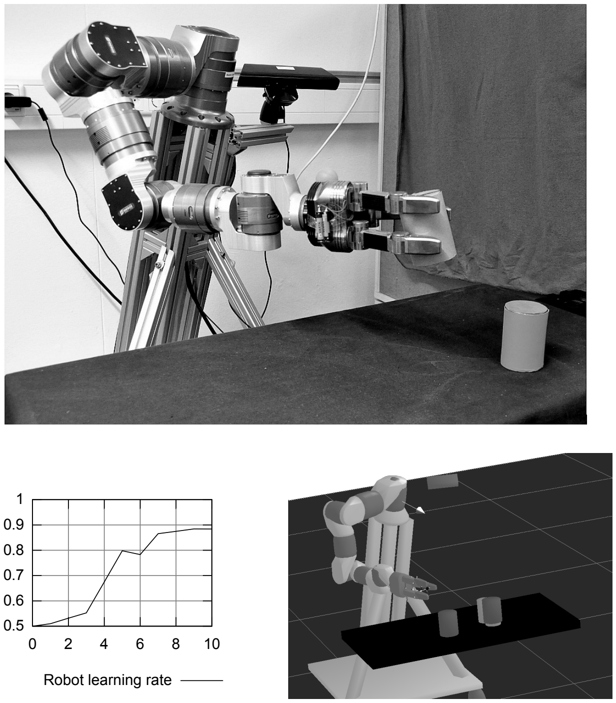


Figure 3: *Experiment 2: Learning on a real world robot.* A qualitative example of learning the symbol *on* with a real-world robot. The robot builds interesting situations (top) according to a sample it generated with active learning (bottom right). After each query it integrates the result to improve its understanding of the predicate. The classification rate is shown in the bottom left. The complete learning trial can be seen in the additional video.

during the start of the *close* learner when using the global criterion. Qualitative experiments suggest that this is an artifact of the small size of the reference set at which the variance is evaluated.

Another interesting investigation is the learning decay of the *upright* learner after 17 samples. The relation is only depending on the angle of the object. Since we use the sine of the angle, we only have a usable feature range of $[-1, 1]$ to learn this relation. The global criterion starts overfitting here, since the gaussian process has a kernel that is too wide to perform better.

Overall the local criterion leads to similar results, although the global criterion outperforms it in all experiments. However, the computing time is much shorter for the local criterion and one might choose it over the global one, if computation time is expensive. The global criterion needs about 15 minutes for a one run, whereas the local criterion finishes it in less than a minute on the same hardware.[1]

4.3 Experiment 2: Real-world robot

To show the possibility of real world applications we performed a complete active learning procedure on a real-world robot as specified above with a human in the loop. The robot is – similar to the first experiment – in the beginning provided with a situation where the relation underlying the symbol holds. It then queries the human teacher by actively manipulating the objects with pick and place actions to generate an interesting situation. After generating the situation the robot senses the real position of the objects and computes the features. It then updates its classifier with the new data. By generating data with object manipulation noise is introduced in the sample generating process, since grasping and putting objects does not always work perfectly on a real robot. The trajectory generator also tries to avoid collisions, hence the real positions differ from the desired. The Kinect perception system introduces another source of noise.

After each query the agent is tested with 5000 simulated situations to test its learning performance. The classification rate of an example trial is shown in Fig. 3.

The experiment shows that learning on an abstract symbolic level can be done by interaction with a human teacher. The method is robust against noise in the sample generation, such as control noise, although the learning rate decreases.

It has been discovered that the rather big embodiment led to problems while generating samples with cylinders very close to each other. The Schunk Dexterous Hand was not able to put the cylinders directly next to each other and the robot therefore tried to generate a very similar sample several times without success. Thus the learning rate decreases after the 7th sample and the classification rate improves less, because the optimal sample could not be learned.

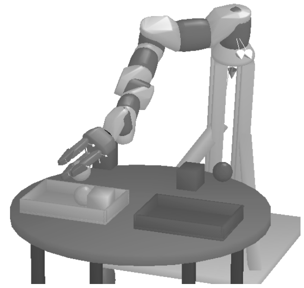


Figure 4: A simulated robot plays with boxes and balls scattered on a table. The task is to put balls in the blue tray and boxes in the red tray.

4.4 Experiment 3: Full-fledged relational RL

We learn relational symbols to use them in relational Reinforcement Learning scenarios. So far relational symbols are hand-crafted and thus not noisy, while learned symbols might include wrong information. In this experiment we tested, whether the learned symbols gives us the ability to plan in complex real-world scenarios.

The task for the robot is to clean up a table with scattered objects. He should put objects of different shape in different trays, until no objects are left on the table (see Fig. 4).

To measure the performance the robot receives a reward for every object within the correct tray. To make faster solutions more profitable the reward is discounted over time, such that every further step decreases the reward.

[1]The used hardware is a standard laptop with current hardware, in detail a current Intel i7, 2.7 GHz dual-core processor and 4 GB RAM.

First, we learned the relation *inside(X, Y)*. This relation is the most important one for the task, since it can show the robot how he performs. The second step is to learn the probabilistic transition model of the action set. For this purpose the robot performs random actions. From the perceived states, which may include the grounded symbol, he learns noisy, indeterministic, deictic rules (NID rules) from the sequences as described in [Pasula *et al.*, 2007]. These rules give a stochastic action outcome based on relational state models. They are used to provide a relational planner with a stochastic model of action outcomes.

Eventually, we use the PRADA planning algorithm [Lang and Toussaint, 2010] to actually plan action sequences and record the reward earned. During planning the robot only has access to the predicted reward based on its belief state. This belief state is possibly wrong, due to the learned symbols it contains. This may lead to false planning strategies. To actually evaluate the robots behavior the real reward is used.

Our experiments are set up as follows. After learning the symbol and the transition model the robot performs 15 actions in a row and after each action the current discounted, cumulated reward is recorded. We performed seven experiments. Three times the symbol was learned with the local criterion and three times it was learned by a random learner. Additionally we performed the same experiment with a perfect oracle as symbol grounding. Each experiment consists of 30 complete runs (whereas a run includes learning of the symbol and transition model).

Fig. 5 shows the results of the experiments. It shows the mean performance of the 30 runs for each experiment and its standard deviation. The figure also shows a more theoretical *optimal behavior*, which is the best possible behavior given that all actions succeed as intended.

It is apparent that learning grounded symbols in an active manner leads to significantly higher rewards in a task using these symbols than random learning. Adding more samples improves both, the active and the random approach, but the gain is bigger when using actively chosen queries.

The difference to the optimal behavior can partly be explained by non optimal behavior of the planner and noisy outcomes of actions (e.g. a ball falling outside a tray), but also partly by non optimal grounding. The behavior with optimal grounding (done by the handcrafted oracle routine) suffers from the noisy control, but outperforms the learned symbol. The optimal behavior shown has no deviation, because of its deterministic action outcomes.

Note that the step like shape of the curves are artifacts from the task, where grasping an object can not lead to reward but only putting things into trays. Thus only every second action can generate reward at best.

5 Conclusions

To enable robots to reason on an abstract level and generalize, symbolic relational representations are a well suited, but handcrafting the groundings is inflexible, time consuming and needs expertise in programming. Teaching robots new symbols in an easy way is thus desirable. We propose an active learning approach to teach robots the groundings of relational symbols with a human teacher. The method is shown to learn spatial relations with few training samples. Hence, it is applicable in interaction with human beings where time is considered expensive. Nevertheless it should be investigated, how well the method scales to more complex tasks and higher dimensional feature spaces.

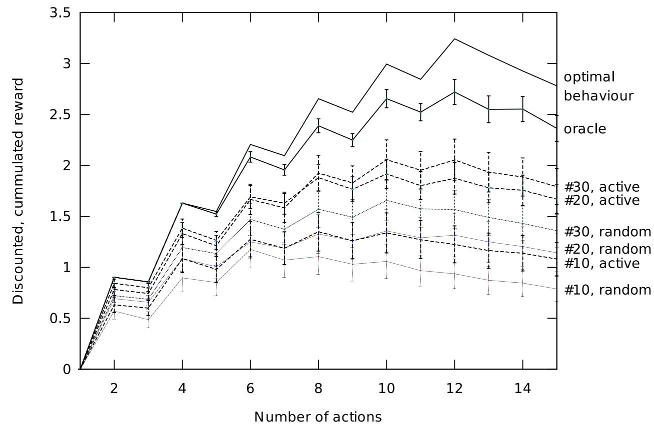


Figure 5: *Experiment 3: Full relational reinforcement learning scenario.* The active learner requires significantly fewer samples to learn grounded relational symbols which permit an autonomous agent to plan for high-reward states.

Future work should investigate methods to ground action symbols, e.g. motor primitives, aiming for a complete grounding of the robots symbols. This would enable non-experts in the field to teach robots for complex tasks. For broader applications of our method, a wider range of feature spaces should be considered. Also developmental approaches could be applied to symbol learning to bootstrap a wider knowledge base with simple symbols. Another direction would be the investigation of possible methods to sample physical situations, as many applications need to reason about real world situations and would benefit from such a method. Such a sampler could also be used for evaluating learning approaches.

Acknowledgments

We would like to thank the anonymous reviewers for their helpful comments. This work is partly funded by the DFG (German Research Foundation) within SPP 1527, *Autonomous Learning*.

References

[Beetz *et al.*, 2010] Michael Beetz, Lorenz Mosenlechner, and Moritz Tenorth. CRAM – A Cognitive Robot Abstract Machine for Everyday Manipulation in Human Environments. In *Proc. of the Int. Conf. on Intelligent Robots and Systems*, pages 1012–1017, 2010.

[Cakmak and Thomaz, 2012] Maya Cakmak and Andrea Thomaz. Designing Robot Learners that Ask Good Questions. In *Proc. of the Int. Conf. on Human-Robot Interaction*, pages 17–24, 2012.

[Cakmak *et al.*, 2010] Maya Cakmak, Crystal Chao, and Andrea Thomaz. Designing Interactions for Robot Active Learners. *Trans. on Autonomous Mental Development*, pages 108–118, 2010.

[Chao *et al.*, 2010] Crystal Chao, Maya Cakmak, and Andrea Thomaz. Transparent Active Learning for Robots. In *Proc. of the Int. Conf. on Human-Robot Interaction*, pages 317–324, 2010.

[Chen and Mooney, 2011] David L. Chen and Raymond J. Mooney. Learning to Interpret Natural Language Navigation Instructions from Observation. In *Proc. of the Nat. Conf. on Artificial Intelligence (AAAI)*, pages 859–865, 2011.

[Cohn *et al.*, 1996] David A. Cohn, Zoubin Ghahramani, and Michael I. Jordan. Active learning with statistical models. *Journal of Artificial Intelligence Research (JAIR)*, 4(1):129–145, 1996.

[Džeroski *et al.*, 2001] Sašo Džeroski, L. de Raedt, and Kurt Driessens. Relational reinforcement learning. *Machine Learning Journal*, 43:7–52, 2001.

[Frank *et al.*, 2008] M. Frank, N. Goodman, and J. Tanenbaum. A Bayesian Framework for cross-situational word-learning. *Advances in Artificial Neural Information Processing Systems*, 20, 2008.

[Harnad, 1990] Stevan Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.

[Katz *et al.*, 2008] Dov Katz, Y. Pyuro, and Oliver Brock. Learning to manipulate articulated objects in unstructured environments using a grounded relational representation. In *Proc. of the Int. Conf. on Robotics: Science and Systems*, pages 254–261, 2008.

[Lang and Toussaint, 2010] Tobias Lang and Marc Toussaint. Planning with noisy probabilistic relational rules. *Journal of Artificial Intelligence Research (JAIR)*, 39:1–49, 2010.

[Lemaignan *et al.*, 2011] Séverin Lemaignan, Raquel Ros, E. Akin Sisbot, Rachid Alami, and Michael Beetz. Grounding the interaction: Anchoring situated discourse in everyday human-robot interaction. *International Journal of Social Robotics*, pages 181–199, 2011.

[Pasula *et al.*, 2007] Hanna M. Pasula, Luke S. Zettlemoyer, and Leslie Pack Kaelbling. Learning symbolic models of stochastic domains. *Journal of Artificial Intelligence Research (JAIR)*, 29:309–352, 2007.

[Rasmussen and Williams, 2006] Carl Rasmussen and Christopher Williams. *Gaussian processes for machine learning*. MIT Press, 2006.

[Salvi *et al.*, 2012] Giampiero Salvi, Luis Montesano, Alexandre Bernardino, and Jose Santos-Victor. Language bootstrapping: Learning word meanings from perception-action association. *Systems, Man, and Cybernetics, Part B: Cybernetics*, pages 660–671, 2012.

[Searle, 1980] John Searle. Minds, brains and programms. *The Behavioral and Brain Science*, 3:417–457, 1980.

[Settles, 2009] Burr Settles. Active learning literature survey. Computer Sciences Technical Report 1648, University of Wisconsin–Madison, 2009.

[Shapiro and Ismail, 2001] Stuart Shapiro and Haythem Ismail. Symbol anchoring in cassie. In *Cognitive Robotics: Papers from the 1998 AAAI Fall Symposium*, pages 136–143, 2001.

[Siskind, 1996] J. M. Siskind. A computational study of cross-sitational techniques for learning word-to-meaning mapping. *Cognition*, 61:39–91, 1996.

[Steels, 2001] Luc Steels. Grounding symbols through evolutionary language games. In A. Cangelosi and D. Parisi, editors, *Simulating the evolution of language*, pages 211–226. 2001.

[Taddeo and Floridi, 2005] Mariarosaria Taddeo and Luciano Floridi. Solving the symbol grounding problem: a critical review of fifteen years of research. *Journal of Experimental and Theoretical Artificial Intelligence*, pages 419–445, 2005.