

Induction of Prototypes in a Robotic Setting using Local Search MDL

Gregory M. Kobele* Jason Riggle* Richard Brooks†
David Friedlander‡ Charles Taylor§ Edward Stabler*

Abstract

Categorizing objects sets the stage for more advanced interactions with the environment. Minimum Description Length learning provides a framework in which to investigate processes by which concept learning might take place. Importantly, the concepts so acquired can be viewed as having a prototype structure - the concepts may apply to one object better than to another. We ground our discussion in a real-world setting - objects to categorize are sensor readings of the behaviours of two mobile robots.

1 Introduction

Natural and artificial agents are confronted with the problem of constructing models of their environments, thereby imposing order on an otherwise chaotic stream of sensory data. An important step in this direction is to recognize when a novel sense datum is sufficiently similar to a previously seen one to make a similar response adaptive. This is the problem of categorization - how does one decide which properties of sense data are relevant, and which are accidental. Obviously, treating all properties as relevant leads to no generalization at all, and treating all properties as accidental leads to maladaptive behaviour (unless the environment is perverse). This problem has been explored in various settings using Minimal Description Length (MDL) concept learning, where the notions of concise encoding of the hypothesis and of efficient encoding of the observed data impose competing pressures (generalizing by minimizing theory size while still being informative about the data), where the measures of theory size and empirical informativeness can be adjusted to fit the domain. We have previously applied this framework both to the clas-

sification of simple robotic behaviour [1], as well as to language evolution [4, 5]. In these previous studies concepts (or languages) were represented as deterministic finite state automata, and percepts were finite strings over a fixed finite alphabet. Our current study extends the previous ones in two ways. First, we enrich the hypothesis space of our agents to include non-deterministic automata. Although deterministic and non-deterministic machines are equivalent in expressive power, enriching the agent's hypothesis space in this manner has repercussions for the results of the learning process. Second, we exploit a particular property of the MDL setup, yielding a cognitively interesting prototype structure of concepts which allows for degrees of membership in a category. One novel aspect of the prototype structure presented here is that the degrees of category membership are a holistic and emergent property of the conceptual space. That is, whether a perception is a "good example" of a particular category depends on what other categories are available. In our setting, concepts are represented as finite automata, and perceptions as strings over a fixed finite alphabet. Given a space of concepts \mathcal{C} , a perception p is a good example of a category $C \in \mathcal{C}$ just in case $p \in L(C)$ and for any other $C' \in \mathcal{C}$ with $p \in L(C')$, the cost of encoding p in C is less than or equal to the cost of encoding p in C' ($enc(p, C) \leq enc(p, C')$).

We present a simple example in which a learner is faced with the task of categorizing the behaviour of two robots, given by means of sensors in the robots' environment. In the learning stage each set of sensor data is labeled as an instance of a concept (e.g. 'random walking', 'wall following', 'chasing', ...). Our robot learner then extracts the salient structure from each training set, possibly generalizing its theory of the training sets to include as yet unseen sensor inputs.

2 Minimum Description Length

The basic idea of the MDL framework is that the best hypothesis to adopt when confronted with a set of data is the one that best navigates between the charyb-

*Department of Linguistics, UCLA. Los Angeles, CA 90095

†Distributed Systems Department, The Pennsylvania State University, State College, PA 16804

‡Informatics Department, The Pennsylvania State University, State College, PA 16804

§Department of Organismic Biology, Evolution and Ecology, UCLA. Los Angeles, CA 90095

dis of theory complexity and the scylla of faithfulness to the data set. In other words, we want the simplest theory with the greatest predictive power. Here we adopt the length of the binary encoding of a machine as our measure of theory complexity, and the cost of specifying a parse of the input (an accepting path through the machine) as our measure of faithfulness to the data. The intuition is that the cost of encoding a string is the amount of information one needs to exactly reconstruct it. This intuition motivates the following definition: as there might be more than one parse of a string through a given machine, we define the cost of encoding a string in a machine as the cost of the *least* path through the machine reading the string.

Formally, we define the cost of the binary encoding of a machine $M = \langle Q, \Sigma, S, F, \delta \rangle$, where Q is a finite set of states, Σ is a finite vocabulary, $S, F \subseteq Q$ are the start and final states, respectively, and $\delta \subseteq Q \times \Sigma_\epsilon \rightarrow 2^Q$ is the transition function, to be

$$\text{size}(M) = \text{size}(\delta) + \text{size}(S) + \text{size}(F)$$

where

$$\text{size}(\delta) = |\delta| (2 \log_2 |Q| + \log_2 (|\Sigma| + 1))$$

and for $Q_0 \subseteq Q$,

$$\text{size}(Q_0) = |Q_0| (\log_2 |Q|)$$

Identifying a parse with a sequence of arcs $\pi = a_1 a_2 \dots a_n$ in M , we define the cost of a parse $\text{cost}(\pi)$ to be the sum of the number of bits needed to specify the path through the machine. Setting z_i to be the number of arcs with the same source as a_i ,

$$\text{cost}(\pi, M) = \sum_{i=1}^n z_i$$

and the cost of encoding a string σ in M is the amount of information needed to recover σ given M ; for Π_σ the set of parses of σ in M ,

$$\text{enc}(\sigma, M) = \min(\{\text{cost}(\pi) | \pi \in \Pi_\sigma\})$$

Given a set of training data T , we first construct the prefix-tree acceptor $P(T)$ for T . Next, we construct the set of machines formed by merging two states in $P(T)$, greedily keeping only those machines M which minimize $\text{size}(M) + \sum_{\sigma \in T} \text{enc}(\sigma, M)$. We continue this process of merging states and discarding machines which are suboptimal until we reach a fixed point.

3 Behavior Recognition with Two Robots

The floor was divided into four rectangular sections and each section was assigned a letter. At any instant, the positions of the two robots can be represented by a pair of letters. The position of the robots in a section was determined by two sensors placed in the room. The sensor setup is shown in figure 1.

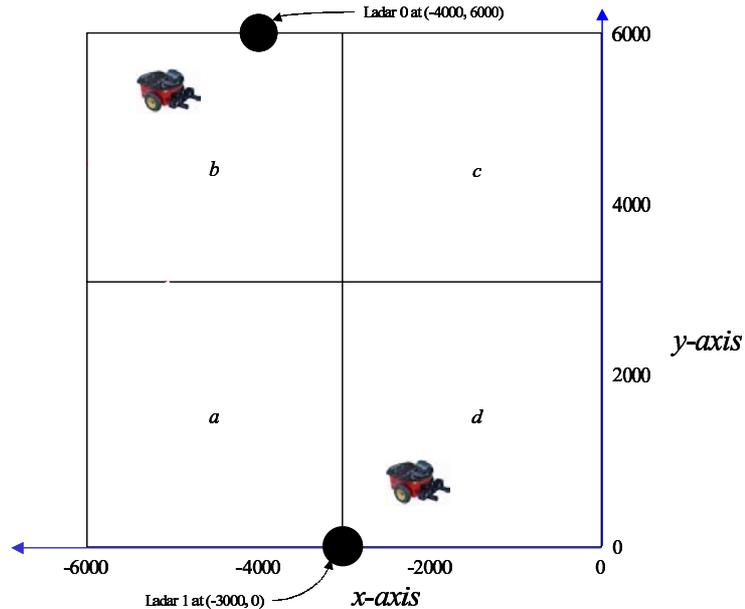


Figure 1: Experimental Setup

Each robotic configuration (i.e. pair of letters indicating the positions of the first and second robots) is taken as a primitive symbol in a larger alphabet (of 16 symbols), into which the sensor readings are ultimately translated. Sensor readings were taken every .2 seconds. As the machines induced by our algorithm for even a five minute time slice were too large to fit readably in the space allotted, we give simplified examples in the next sections.

4 Prototypes

Prototype theory [3] is a theory of concepts which is intended to account for behavioural asymmetries in identifying objects as falling under particular concepts, as well as regularities in ‘prototypicality judgements’ across individuals. A model of this type of

data, should minimally provide an ordering relation across entities participating in the concept. Osherson and Smith [2] challenged the sufficiency of this minimal criterion of adequacy. They argued that fuzzy set theory [6] does not provide an adequate model of the data, even though fuzzy set membership does provide just such an ordering over individuals. They point out that the theory of concepts should be able to account for the compositional nature of concepts - that is, that some concepts are formed from others. They claim that concept compositionality is not faithfully reflected in Zadeh’s fuzzy set operators. As a simple example they chose the concept *striped apple*, which they took to be composed of the (more) atomic concepts of *striped* and *apple* via concept conjunction. On the assumption that a prototypical striped apple is neither a prototypical striped thing nor a prototypical apple, this fact is not well-modeled using Zadeh’s fuzzy set intersection operator, which identifies the degree of membership of an object in the conjunction of two fuzzy sets with the least degree of membership of that object in the conjuncts.

In the model presented herein, the cost of encoding a percept into a machine yields an ordering across participants of a concept, with p being a more prototypical M than p' iff $enc(p, M) < enc(p', M)$. If we take the standard operations of machine intersection (cross-product of machines followed by removal of dead states and arcs) and machine union (creation of a new start state, and empty transitions to the start states of each concept) to model the binary conceptual connectives ‘and’ and ‘or’, we see that in the ‘or’ case our model and the fuzzy set model (in which an element is in the union of two sets to the greater of the degrees in which it is in each of the individual sets) coincide (recall that in the fuzzy set model, the higher the degree of membership, the more prototypical, and in our model, the lower the cost of the encoding, the more prototypical). As shown below (in figures 2, 3 and 4), the removal of dead states (and, more importantly, of the arcs to them) after machine intersection can increase the goodness of fit of the conjoined concept to the instance. This allows for a prototypical striped apple to be a better example of the concept *striped apple* than of either *striped* or *apple*.

Using the floor layout described in figure 1, figure 2 represents the concept of a single robot remaining arbitrarily long in section a , then moving to b for two sensor readings, and finally possibly going to c and staying there.

Figure 3 represents the concept of a single robot remaining in a for two sensor readings, then possibly

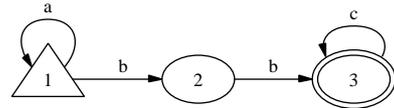


Figure 2: machine M_1 recognizes $L_1 = a^*bbc^*$. The strings with the cheapest encoding (and highest prototypicality) are those in bbc^* .

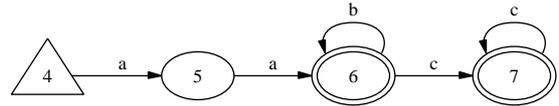


Figure 3: machine M_2 recognizes $L_2 = aa b^*c^*$. The most prototypical strings are in aac^* .

moving to b and remaining there arbitrarily long, and finally possibly going to c and staying there.

Finally, figure 4 represents the concept which is the conjunction of the two previous ones; namely, that of a single robot remaining in a for two sensor readings, then moving to b and staying there a like amount of time, and then possibly moving to c and staying there indefinitely. The observation $aabb$, which is of the robot staying in a for two turns before moving to b and staying there for two turns, participates in all three concepts, although it is cheapest to describe (and thus most prototypical) in M_3 (see figure 5) as desired.

5 Summary

We have presented a model of categorization which results in a natural ordering between instances of a concept, as well as a way to determine which concept best fits a new instance. This latter point might be extremely valuable in practice, especially as having a way to discriminate between concepts which apply to the given instance may allow for a more adaptive response. The model also seems able to surmount some intuitive difficulties previous models of concepts had in regard to the compositional nature of conceptual structure.

Acknowledgements

This work was supported by the UCLA Center for Embedded Network Sensors, the Defense Advance Research Projects Agency (DARPA), administered by

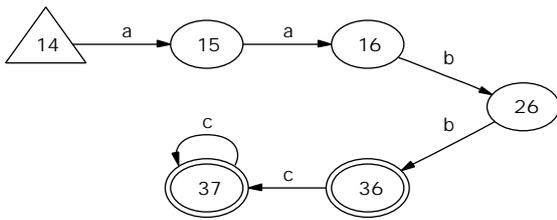


Figure 4: machine $M_3 = M_1 \times M_2$ recognizes $L_3 = L_1 \cap L_2 = aabbc^*$. The most prototypical strings are those in $aabbc^*$.

	Encoding
M_1	2
M_2	2
M_3	0

Figure 5: A comparison between the costs of encoding the string $aabb$ in each machine.

the Army Research Office under Emergent Surveillance Plexus MURI Award No. DAAD19-01-1-0504, and DARPA MURI award administered by the US Airforce No. F49620-01-1-0361. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

References

- [1] Yoosook Lee, Jason Riggle, Travis C. Collier, Edward Stabler, and Charles E. Taylor. Adaptive communication among collaborative agents: Preliminary results from symbol grounding. In Masanori Sugisaka and Hiroshi Tanaka, editors, *Proceedings of the Eighth International Symposium on Artificial Life and Robotics*, pages 149–155, 2003.
- [2] D. N. Osherson and E. E. Smith. On the adequacy of prototype theory as a theory of concepts. *Cognition*, 11:237–262, 1981.
- [3] E. Rosch. On the internal structure of perceptual and semantic categories. In T. E. Moore, editor, *Cognitive development and the acquisition of language*. Academic Press, New York, 1973.
- [4] Tracy Teal, Daniel Albro, Edward Stabler, and Charles E. Taylor. Compression and adaptation. In J. Nicoud D. Floreano and F. Mondada, editors, *ECAL99*, pages 709–719, Berlin, 1999. Springer-Verlag.

- [5] Tracy K. Teal and Charles E. Taylor. Effects of acquisition on language evolution. *Artificial Life*, 6(2):129–143, Spring 2000.
- [6] L. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.