

Adaptive Communication among Collaborative Agents: Preliminary Results with Symbol Grounding

Yoosook Lee[†] Jason Riggle[§] Travis C. Collier[†] Edward Stabler[§]
Charles E. Taylor^{†*}

Abstract

Communication among adaptive agents can be framed as language acquisition and broken down into three problems; *symbol grounding*, *language learning*, and *language evolution*. We propose that this view clarifies many of the difficulties framing issues of collaboration and self-organization. Additionally, we demonstrate simple classification systems that can provide the first step in grounding real-world data and provide general schema for constructing other such systems. The first system classifies auditory input from frog calls and is presented as a model of grounding objects. The second system uses the minimum description length framework to distinguish patterns of robot movement as a model of grounding actions.

1 Introduction

1.1 The Vision and Approach

Adaptive collaboration holds great promise for artificial life and engineering generally. At present there is no good theory of self-organization and collaboration, though the benefits are obvious. These include concurrent specialization by different individuals; distributed presence (for triangulation or “extending the eyes and ears of a small party”); robustness of individuals or functions that can be assumed by others; ability to avoid high risk or high cost behaviors until needed, among others.

The long term goal of our research is to produce a collection of heterogeneous robots that learn about their environment, communicate with one another about it, affect the environment in ways that take advantage of one another’s special abilities and circumstances, and report to human observers.

The work described here is directed toward creation of robots, or more abstractly agents, which can learn to bind symbols to sensory patterns from

their environment and to evolve a language that will enable them to use those symbols for meaningful communication among themselves. This language will also provide a logic to facilitate reasoning.

1.2 Language at the Core

At this point the principal challenges are to further develop the cognitive and linguistic features of the system. It is desirable that agents be able to acquire their own language. First, an acquired language can evolve over time to adapt its structure to various types of noise and different types of information to be transmitted, thus providing a truly flexible communications channel. Also, agents with heterogeneous sensor modalities can map the same symbol to an environmental stimulus despite having very different internal representations. This requires agents to acquire their own symbol-to-meaning mappings, based on how their own observations correlate with the symbols that other agents transmit. Finally, language also provides a logical manipulation facility for cognitive reasoning. Our approach has been explained in more detail elsewhere¹.

*taylor@biology.ucla.edu

[†]Department of Organismic Biology, Ecology, and Evolution, University of California Los Angeles

[§]Department of Linguistics, University of California Los Angeles

1.3 Language Acquisition

Language acquisition provides a well explored framework for many of the problems in designing collaborative agents. From a linguistic perspective, the primary problems are: symbol grounding, language expression and learnability, and language evolution.

The problem of *symbol grounding* is, given a finite sequence of sensory impressions, each with potentially infinite detail and accompanied by a sentence, how does one come to associate particular lexical items intended meanings²? Symbol grounding, in the most general case, is known to be intractable. However, Siskind³ and others have shown that when there is a consistent shared hierarchy of salience, then computational systems can overcome the grounding problem in quite general circumstances.

Language learning is the identification of a possibly infinite language and its syntactic rules from a finite sequence of examples. For a long time, reasonably expressive languages were considered unlearnable⁴; however, recent progress by Angluin⁵, extended by Kanazawa⁶, Denis⁷, and Stabler^{8, 9}, has demonstrated that a class of quite expressively powerful languages can be learned.

The third problem, *language evolution*, involves determining how the language used by a collection of agents will change over time. This aspect is closely related to just what information needs to be expressed and to the kinds and quantities of noise the agents have to deal with. Parameters such as the maximum amount of communication channel noise that can exist before the agents lose their ability to communicate in a common language have been determined for simple regimes¹⁰. Doing such analysis for the system we envision may not be analytically tractable, but empirical studies informed by the theoretical results for simpler systems are possible.

In this paper we will focus our discussion on the the first of these language acquisition problems – symbol grounding.

1.4 Symbol Grounding

In symbolic systems, transfer of information is accomplished by associating symbols with words, gestures, or other behaviors that can be sensed by oth-

ers. Establishing this association between a symbol and the information it conveys is called *grounding*. In an adaptive system, the meaning of a symbol might vary, and so a language learner needs to somehow ground its linguistic experience in the cognitive domain, establishing the semantic values of the symbols.

Grounding, the learning of culturally common semantic values on the basis of similar experience, also allows for language change. Like other instances of phenotypic plasticity, this can allow a single organism to adapt and survive where a fixed system might fail. Humans can quite quickly adapt to widely different sorts of vocal tracts and other synthesizers in spite of vastly different acoustic properties, and even when vocal communication becomes impossible, there are often other means. Robots can have similar capabilities. Note in particular that in the situation where two robots already have had a common language and previous communication so that each has a model of the other, then if a change in language is needed, grounding the changed language can be very quick and efficient.

Various approaches to the roles of perception and language adaptation in grounding problems have been explored^{11, 12, 13}. Here we focus on ideas in simple acoustic classification and the *minimal description length* (MDL) framework^{14, 15, 16}.

2 Grounding Objects

The first task any symbolic communication system must solve is how to classify non-symbolic sensory input. This is a prerequisite for further processing functions, such as localization. For the communication that allows collaboration to occur, a classification system is needed to bind actual perceptual information into symbolic tokens that can carry the semantic meaning.

Our model for this step has been an aural recognizer that quickly discriminates between different species of frog calls with reasonably low error. Frog calls provide a good model because as they are typically simple, short, and well delimited. As such, they are useful in a variety of model situations, including localization¹⁷. Besides their intrinsic interest to biologists, frogs are often a bellwether of ecological disturbance such as pollution

and climate change, so that, monitoring species distributions and abundances could provide an early warning of environmental problems.

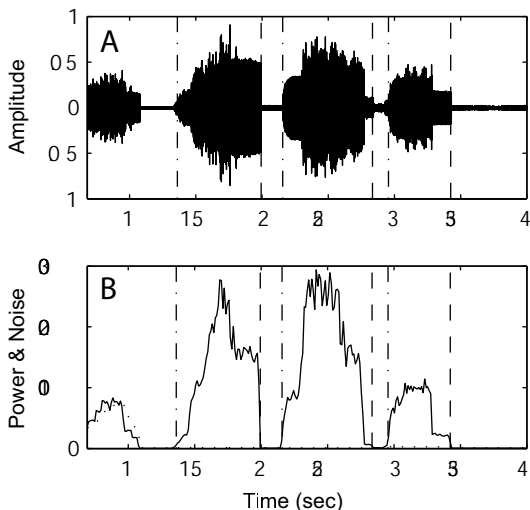


Figure 1: Signal detection of frog calls. Detected starting points ($- \cdot -$) and ending points ($- - -$). **A.** Amplitude of input source. **B.** Power and estimated noise ($\cdot \cdot \cdot$) derived from amplitude data.

Our system works as follows. Given raw auditory input the agent starts off by determining the beginning and ending points of potential *signals* as shown in Figure 1A. This is accomplished by keeping a running estimate of noise and looking for increases (or decreases) in the signal-to-noise ratio as depicted in Figure 1B. When a signal has been detected, the spectrogram for the signal is computed and compared to spectrograms of previously identified sounds. This signal spectrogram comparison is simply the maximum cross correlation coefficient which has been shown to properly correlate to the probability that two samples are identical except for additive white-noise¹⁸. If the new signal does not correlate well with any previously identified sounds, it is considered to be a new sound, and is saved so that additional signals can be compared to it as well as to the previously identified ones.

Figure 2 shows the first results from real data (the published recorded calls of frogs and toads in French Guyana¹⁹); 1507 potential signals from 62 different frog species. Using 0.7 as the similarity threshold for matching and 0.65 as the threshold

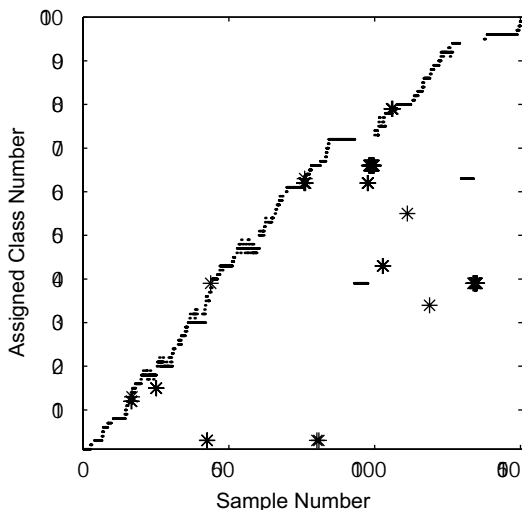


Figure 2: Classification performance from frog calls. Correct classification (\cdot); Incorrect classification ($*$). Each new signal is either assigned to an existing class or creates a new class.

below which the signal is considered new, we failed to classify 15% of the signals and misclassified less than 5%.

This general strategy of detecting potential signals, transforming the signal data to extract relevant features for comparison, and then applying a statistical comparison to assign the signal to a class of known signals (or else create a new class of known signals) is easily extended. More sophisticated techniques are possible, and we are currently collaborating with engineers and computer scientists to explore pitch, duration, power-spectra, and other measurements. Multiple metric analysis opens up the possibility of doing hierarchical classification, including subclasses that correspond to adjectives.

Although far from perfect, this simple method works much better than one would expect after reading the very discouraging literature on related tasks such as speech-recognition²⁰. Interesting patterns that can potentially distinguish real biological hypothesis are discernible even at this stage, although clearly more work needs to be done. We are currently working with herpetologists to explore strength of species relatedness, environmental conditions, and inter-specific competition for bandwidth in determining the similarity between differ-

ent frog species calls.

3 Grounding Events

Sequences of symbolic data, either produced by object grounding or directly from sensors, can also be grounded with linguistic inputs to produce the equivalent of verbs and adverbs. Our work in this direction is focused on analyzing behavioral data from animals and robots using the MDL as a basis for generalization.

3.1 Minimal Description Length

The MDL framework^{14, 15, 16} is essentially an implementation of standard Bayesian induction, where the priors are set by the representational complexities of the hypotheses. The basic idea is that the learner prefers the simplest description of the evidence, where the measure of simplicity is given by the representational complexity of the hypothesis together with its encoding of the evidence. A hypothesis that specifies a sufficiently general tendency in the data will make the encoding of the data smaller, for an overall improvement in simplicity.

Many different learning models can be represented in this framework, and they have proven useful already in theories of perception²¹ and language learning^{22, 23, 24, 25}. We have deployed these models in a few of preliminary domains: automata induction²⁶, and recently in identifying wall-following and other robotic behavior.

Finding an absolutely minimal description is impossible and unnecessary. The best compression of a set of finite strings is not, in general, a computable function²⁷. Moreover, for simple, finite strings, we are often interested in differences smaller than the constant bound distinguishing different universal machines²⁸. For our purposes, a simple and efficient metric of hypothesis complexity will be most useful in guiding generalization.

3.2 A Simple Example of MDL

A simple example is presented in Figure 3 which illustrates sensors distributed in a room. A robot wandering around the room activates the sensor in the region it occupies. The following stimulus sets

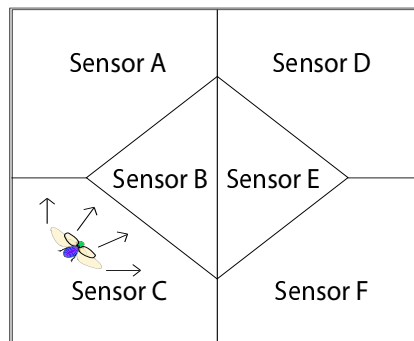


Figure 3: A robot wandering around a room where array of sensors are distributed.

include elements that represents the movement of a robot at each time, kindly provided to us by Brooks and Friedlander.

$$A = \left\{ \begin{aligned} &[a, a, e, f, f, f, d, d, f, c, a, a], \\ &[c, c, f, f, c, b, b, b, b, f, c, a], \\ &[a, b, b, c, f, f, e, a, a, d, f, e] \end{aligned} \right\}$$

$$B = \left\{ \begin{aligned} &[c, d, f, a, c, d, f, a, c, d, f, a], \\ &[c, d, f, a, c, d, f, a, c, d, f, f, f], \\ &[c, d, f, a, c, d, f, a] \end{aligned} \right\}$$

The question is, can we characterize the difference between random and patterned behavior of this robot?

We will represent our hypotheses as deterministic finite state automata(FSA). Even with this restricted representation scheme finding a globally optimal encoding is not feasible because the range of possible hypotheses grows exponentially with the amount of data in the sample.

At one extreme we can generate a machine which accepts only previously seen examples as instances of the concept. This can be encoded as a deterministic FSA that accepts exactly the examples seen[Figure 4 *A, B*]. At the other extreme we can encode concepts with a universal acceptor which simply accepts any string at all [Figure 4 *A'*].

The goal here is to find the optimal point between these two extremes that generalizes beyond the observed data but not too far. To this end we propose the local search algorithm presented in Table 1, which finds an approximately minimum description length. It is similar, but not identical, to that used earlier by Teal *et al.*²⁹.

rectly distinguishes the random behavior in data set A from the patterned behavior in data set B . Finally we should note that while this metric of description length was effective in discriminating simple patterns in the data for robot movement behavior, different domains will require different metrics of generalization.

4 Discussion and Conclusion

The framework provided by viewing collaboration of heterogeneous agents as language acquisition provides a well explored theoretical basis for many of the most vexing problems. As a first step in implementing such a language acquisition system, we have bound real-world sensor data to symbols using both a classification system for object grounding and an MDL approach for event grounding.

Within the scope of auditory data from frog calls a simple classification system has been implemented and provides a general model that can be easily extended. The combination of a relatively easy data source combined with the fact that the linguistic grounding process can tolerate some misclassification errors motivates more sophisticated work on good, but not perfect, classifiers.

The MDL strategy presented above is effective for distinguishing among some simple patterns of sensory data corresponding to different types of events. The next logical step in attempting to ground events is to develop strategies for more complex events and noisier data. To this end we have begun preliminary investigations using MDL techniques on ethograms.

Originally devised by naturalists studying animal behavior, ethograms catalog the behavioral events (or specific action patterns) that animals use during different behavioral states or contexts. Since ethograms can be encoded as series of behavioral events, it seems plausible to extract a generalized pattern of behavior of a given animal from its ethogram using MDL method.

Our initial examinations of marmot ethograms (obtained from Dr. David Blumstein at UCLA) using the MDL method presented above did not yield any patterns that could be encoded with deterministic finite automata other than a universal acceptor. This serves to illustrate the crucial relationship between the representation scheme used to formu-

late the hypotheses and the concepts that are the target of generalization. For future investigation of ethograms and other noisy patterns probabilistic nondeterministic finite automata may prove more effective. We are currently exploring this possibility.

Acknowledgments

We thank Richard Brooks and David Friedlander for discussion and use of their robot data, Hanbiao Wang provided helpful suggestions for analysis of the frog data. This work was supported by the UCLA Center for Embedded Network Sensors, the Defense Advance Research Projects Agency (DARPA), administered by 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] K. Wee, T. Collier, G. Kobele, E. Stabler, and C. Taylor. Natural language interface to an intrusion detection system. In *Proceedings, International Conference on Control, Automation and Systems*. ICCAS, 2001.
- [2] S. Harnad. The symbol grounding problem. *Physica D*, 42:335–346, 1990.
- [3] J. M. Siskind. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61(1-2):39–91, 1996.
- [4] E. M. Gold. Language identification in the limit. *Information and Control*, 10:447–474, 1967.
- [5] D. Angluin. Inference of reversible languages. *Journal of the Association for Computing Machinery*, 29:741–765, 1982.
- [6] M. Kanazawa. Identification in the limit of categorial grammars. *Journal of Logic, Language, and Information*, 5:115–155, 1996.
- [7] F. Denis. Learning regular languages from simple positive examples. *Machine Learning*, 44:37–66, 2001.
- [8] E. P. Stabler. Acquiring grammars with movement. *Syntax*, 1:72–97, 1998.
- [9] E. P. Stabler. Identifying minimalist languages from dependency structures. Presentation available at <http://www.clsp.jhu.edu/seminars/abstracts/S2002/stabler.shtml>, 2002.
- [10] N. Komarova, P. Niyogi, and M. Nowak. Evolutionary dynamics of grammar acquisition. *Journal of Theoretical Biology*, 209(1):43–59, 2001.
- [11] L. Steels. Self-organising vocabularies. *IEEE Intelligent systems*, pages 16–23, 2001.

- [12] S. Kirby. Spontaneous evolution of linguistic structure: an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation*, 5:102–110, 2001.
- [13] M. Oliphant. The dilemma of saussurean communication. *BioSystems*, 37:31–38, 1996.
- [14] V. N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, NY, 2000.
- [15] P. Vitányi and M. Li. On prediction by data compression. In *Proceedings of the 9th European Conference on Machine Learning*, Lecture Notes in Artificial Intelligence, Vol. 1224, pages 14–30, 1997.
- [16] P. Vitányi and M. Li. Minimum description length induction, Bayesianism, and Kolmogorov complexity. *IEEE Transactions on Information Theory*, IT-46:446–464, 2000.
- [17] H. Wang, J. Elson, L. Girod, D. Estrin, and K. Yao. Target classification and localization in a habitat-monitoring application. submitted to ICASSP 2003, Hong Kong, 2003.
- [18] H. Wang, L. Yip, D. Maniezzo, J. C. Chen, R. E. Hudson, J. Elson, and K. Yao. A wireless time-synchronized cots sensor platform part ii: Applications to beamforming. In *In Proceedings of the IEEE CAS Workshop on Wireless Communications and Networking, Pasadena, California, September 5-6, 2002*. IEEE, 2002.
- [19] C. Marty and P. Gaucher. Guide sonore des amphibiens anoures de guyane. Compact-Disc, 2002.
- [20] L. Rabiner and B.-H. Juang. *Fundamentals of Speech Recognition*. Englewood Cliffs NJ: PTR Prentice Hall (Signal Processing Series), 1993. General Intro : ISBN 0-13-015157-2.
- [21] D. Mumford. Pattern theory: a unifying perspective. In D. C. Knill and W. Richards, editors, *Perception as Bayesian Inference*. CUP, NY, 1996.
- [22] M. R. Brent and T. A. Cartwright. Lexical categorization: Fitting template grammars by incremental MDL optimization. In L. Micla and C. de la Higuera, editors, *Grammatical Inference: Learning Syntax from Sentences*, pages 84–94. Springer, NY, 1996.
- [23] J. Rissanen and E. Ristad. Language acquisition in the MDL framework. In E. Ristad, editor, *Language Computations*. American Mathematical Society, Philadelphia, 1994.
- [24] P. Grünwald. A minimum description length approach to grammar inference. In G. S. S. Wermter, E. Riloff, editor, *Symbolic, Connectionist and Statistical Approaches to Learning for Natural Language Processing*, LNCS #1040. Springer-Verlag, Berlin, 1996.
- [25] J. Goldsmith. Unsupervised learning of the morphology of a natural language. *CL*, 27(2):153–198, 2001.
- [26] T. K. Teal, D. Albro, E. Stabler, and C. E. Taylor. Compression and adaptation. In *European Conference on Artificial Life, ECAL'99*, 1999.
- [27] A. Kolmogorov. Three approaches to the definition of the concept 'amount of information'. *Selected Translations in Mathematical Statistics and Probability*, 7:293–302, 1968.
- [28] A. N. Soklakov. Complexity analysis for algorithmically simple strings. *arXiv:cs.LG*, 0009001, 2001.
- [29] T. K. Teal and C. E. Taylor. Effects of acquisition on language evolution. *Artificial Life*, 6(2):129–143, Spring 2000. als titled as: Effects of Compression on Language Evolution.