Language-Learning Determines All Learning

Leona F. Fass

Abstract

We describe results in formal language-learning whereby a learner observes and experiments, dynamically interacting with its environment to acquire a model of linguistic knowledge. While techniques for achieving such learning are affected by properties of the knowledge domain and the learner's environment, we find they are adaptable to other learning problems.

Introduction

We first investigated learning with an emphasis on the acquisition of linguistic knowledge and then we considered a more general theory of "learnability". We found that critical issues arising in problems of language-learning are issues that must be confronted when many other learning problems are to be solved. These include devising techniques for representing and conveying knowledge; selecting illustrative examples and processing observed data; and establishing criteria for acceptance of a learned result. Each of these issues may be affected by the domain in which the body of knowledge lies, and the environment in which the "learner" resides. In fact, through our language-learning research we have concluded that such grounding can *define* a learning technique, as well as defining the resultant model with which specified knowledge may be acquired.

We now describe the techniques we developed, and the results we obtained, in the specific area of formal language learning. Then we discuss the potential applicability of our techniques to all processes of learning.

Our "Grounded" Language-Learning

We initially took a formal, foundational approach to language-learning, algebraically analyzing the knowledge in a linguistic domain. Generalizing some classical machine theory (Myhill 1957, Nerode 1959) we sought generative and recognitive processors to characterize the formalized linguistic knowledge. *If* a learner could construct such a finite processor to model the entire language, we would conclude that the language had been *learned.* All of the knowledge in the linguistic domain would be acquired, through acquisition of a finite model.

It had been suggested by (Levy and Joshi 1978) that language might be conveyed structurally so we imposed a "structural bias" on the linguistic domain (representing sentences as skeletal derivation trees). We then found that a learning process could *exploit* the formal representation and the domain constraints, to obtain both generative and recognitive models. *Dynamic interaction* between a "learner" and its environment (the language domain) could produce the desired or "best possible" result. In our original language-learning work we even found that a learner could interact directly with the language domain and determine *minimal* language models (Fass 1996).

A situated learner, beginning with *no* language model, could observe external information to determine (as congruence classes of substructures within the knowledge domain) those disjoint categories of linguistic objects that a characterizing language model would *need.* (Some syntactic objects might even be determined from semantic observations and segmentation, as in infant learning.) Experimentation with combinations of objects could then determine "correct" linguistic groupings: these would be represented by syntactic rules. Once observations and experiments were completed, the rules and categories of objects would define the learned linguistic model. (If new domain elements were added, the set of categories and set of rules could *monotonically* expand.) Tagging of (skeletal) language structures from the domain could be effected, if desired, once the syntactic model was obtained (Fass 1998a).

We found that *if* a finite model of a language domain exists, then necessary experiments *will* be completed and the model found, effectively. Furthermore , if a minimal model is known, a learner can infer it from specific observations of the domain. But if *no* such model is known then at any time experimentation is halted, the monotonic learning process will identify a potential language model, finitely and approximately. This representation may meet

established learning criteria and be accepted as "best possible" or "good enough".

Conclusions and Adaptations to Other Learning Problems

We could obtain these satisfying results in the language-learning case *only* because of the relationship established between the learner's linguistic environment and the language's learnable model. (Categories of domain elements corresponded to components of the model. Representatives of these categories provided sufficient selected information from which the model may be inferred.) This relationship existed only because we imposed domain constraints and a structural bias that could *force* the discovery of a finite learnable result. By imposing domain structure, we found we had imposed finite realizability, representation and learnability.

Although we developed a learning process for the linguistic domain, precisely the same procedures can apply in any other environment where structure and finite representation might exist or be imposed (Fass 1998b). Learning would still be effected by discovering disjoint knowledge categories (again, congruence classes of structures) and constructive rules. A situated learner, interacting with the knowledge domain by experimenting with representatives of these categories, might construct a finite, minimal or "best possible" result. While discretely realizing domain knowledge, the process developed for language-learning thus determines *all* learning.

Selected References

Bienkowski, M., 1998. "A Reader's Guide to Agent Literacy", *SIGART Bulletin*, Vol 9 No 2 (Fall 1998), 23 -29.

Fass, L. F. , 1996. "Discovering a 'Minimal' Language Model", *Proc of the Ninth Florida AI Research Symposium*, Key West, May 1996, 349 - 353.

Fass, L.F., 1998a. "Formal Learning Constraints and Discourse Processing", invited for participation in *AAAI Spring Symposium on Applying Machine Learning to Discourse Processing*, (3pages), Stanford, 1998.

Fass, L.F., 1998b. "Learning (Language) within a Context", *Proc. of the Fourth Joint Conf on Information Sciences / Third Internatl Conf on Computational Intelligence and Neuroscience*, Research Triangle Park, NC, October 1998, Vol II, 56 - 59.

Gold, E.M., 1967. "Language Identification in the Limit", *Inf. Contr.*, Vol 10 (1967), 447-474.

Levy, L.S. and A.K. Joshi, 1978. "Skeletal Structural Descriptions", *Inf. Contr.*, Vol 39 (1978) 192-211.

Myhill, J., 1957. "Finite Automata and the Representation of Events", *WADC Tech Rept*, 57-624, Wright-Patterson AFB, 1957.

Nerode, A., 1959. "Linear Automaton Transformations", *Proc. AMS*, IX (1959), 541 - 544.

Valiant, L. , 1984. "A Theory of the Learnable", *CACM*, Vol 27 (1984), 1134 - 1142.

Leona F. Fass received a B.S. in Mathematics and Science Education from Cornell University and an M.S.E. and Ph.D. in Computer and Information Science from the University of Pennsylvania. Prior to obtaining her Ph.D. she held research, administrative and/or teaching positions at Penn and Temple University. Since then she has been on the faculties of the University of California, Georgetown University, and the Naval Postgraduate School. Her research primarily has focused on language structure and processing; knowledge acquisition; and the general interactions of logic, language and computation. She has had particular interest in inductive inference processes, and applications/adaptations of inference results to the practical domain. She may be reached at

Mailing address: P.O. Box 2914
Carmel CA 93921
lff4@cornell.edu