Language Learning in Large Parameter Spaces

Karen T. Kohl
MIT Artificial Intelligence Laboratory
545 Technology Square, Room 809
Cambridge, MA 02143
ktkohl@ai.mit.edu

Introduction

Various theories of linguistics have proposed that the differences among natural languages can be parameterized. Certainly syntactic theories such as Principles and Parameters (Chomsky, 1981) assume the existence of such parameters. Along with the problem of defining parameters, we need to address the problem of a child's acquisition of the settings of these parameters.

Several algorithms for parameter setting have been proposed and examined on small spaces. Unless we have a realistic space to study, we cannot fully understand the predictions of these algorithms. Having an implemented computational model of these algorithms is important for studying them at greater depths. This study examines one such parameter-setting algorithm in realistic spaces.

The TLA

Gibson and Wexler (Gibson and Wexler, 1994) propose the Triggering Learning Algorithm, or TLA, for binary-valued parameters. If the learner hears a trigger, or a sentence which she cannot analyze under the current parameter settings, she randomly selects one parameter and changes its value if the new value allows her to analyze the sentence. Niyogi and Berwick (Niyogi and Berwick, 1996) have characterized the TLA as a local hill-climbing search algorithm with memoryless learning.

Gibson and Wexler's proposal studied a space with only three parameters. They found that certain target languages were not learnable if the learner started in certain states, called local maxima. One proposed solution was to start with a default value for one parameter. Another proposal was that the local maxima would disappear with a larger, more realistic parameter space.

A 12-Parameter Space

To examine potential solutions to the local maxima problem, we used Stefano Bertolo's implementation of the TLA. His system added nine syntactic parameters to Gibson and Wexler's original three parameters. The twelve parameters involved position and. We added several enhancements allowing users to select and to add parameters.

We found that the likelihood of finding local maxima in a space increased with the size of the space. As the spaces grew more realistic, more parameter interaction resulted in more local maxima. Therefore, the problem of local maxima is not specific to Gibson and Wexler's three-parameter space and does not disappear as the space becomes more realistic.

Since we found that local maxima were common in this large twelve-parameter space, we considered several solutions. When looking at default settings of parameters, we found no initial setting of all parameters that would guarantee the learnability of all target languages.

Next we tried to find default settings of a few parameters that would make most target languages learnable. If we could find parameter settings that patterned together in the unlearnable languages, then we could predict and try to verify that these kinds of languages are unattested in natural language. However, we found that the best default settings still predicted several known natural languages to be unlearnable.

Conclusion

The problem of local maxima is very real for the TLA. More parameters means that there will be more interaction. The solutions of using default settings of parameters made the wrong predictions.

Although we did not find a solution to the problem of local maxima in the TLA, we did show how to study the algorithm in depth. Without this implementation, we could not have been sure of the extent of the problem of local maxima. Any theory of the acquisition of parameter setting should be studied closely with large parameter spaces.

References

Broihier, K. 1997. Case Studies in Language Learnability. Ph.D. diss., Dept. of Brain and Cognitive Science, MIT.

Chomsky, N. 1981. *Lectures on Government and Binding.* Dordrecht: Foris.

Gibson, E., and K. Wexler. 1994. Triggers. *Linguistic Inquiry* 25:407-454.

Niyogi, P., and R. Berwick. 1994. Learning from Triggers. *Linguistic Inquiry* 27:605-622.

Niyogi, P., and R. Berwick. 1996. A Language Learning Model for Finite Parameter Spaces. *Cognition* 61:162-193.