“Real-time” Language Use by “Real” Agents

Penelope Sibun
P. O. Box 620527
Woodside CA 94062
sibun@parc.xerox.com

People do not speak in the well-formed sentences that we find in written text. In fact, a careful transcription of spoken language will reveal that it is generally composed of short fragments of text that are strung together. While hierarchical sentence structure can often be imposed on such texts, an alternative analysis that takes the fragment as the basic unit of language accounts for most linguistic phenomena, and has the advantage of constituting a simpler explanation. In particular, a fragment-based analysis is well-suited to a model of incremental text production.

In my thesis (Sibun 1991), I presented Salix, a program for incremental text generation. Salix produces texts about things that are obviously structured, and uses the structure of the subject matter to structure the text it produces. The types of fragments that Salix produces, and the strategies it uses for choosing what text to generate next, were created by examining how people generate texts in similar situations. Salix’s major limitations lie in the constraints imposed by the initial selection of situations in which it could produce text. Salix has so far been limited largely to monologues about richly-structured objects such as families and house layouts, about which it has readily available all the information it will use.

I am working to broaden Salix’s horizons in two directions simultaneously. First, it will be a greater test of Salix’s ability to always find something to say next if it is learning about its subject matter at the same time at which it is describing it. For example, suppose Salix takes its knowledge of how to talk about a building’s layout and explores a real building that it has not seen before. In this case, Salix needs a way to extract information from the environment that it can map onto its strategies for producing text.

Second, I would like to expand Salix’s environment to include other agents who would also be generating text of their own. Salix’s context includes information about what has already been said. Heretofore, the only speaker of interest has been Salix itself; in order to make an intelligent choice of what to say next in a multiagent environment, Salix needs to be able to incorporate into its context what other agents have said as well. An obvious way of combining these two goals is to join Salix’s language capabilities with those of a mobile robot with vision.

A real-time environment that is changing both physically and by virtue of interaction with other agents is a demanding one. I believe that for an artificial agent to have any hope of using language in such a context it must be able to make simple choices quickly and continuously about what to say next. These choices will always be highly constrained by the changing context, and will each result in a short segment of appropriate text. The emphasis in such text generation will not be on producing well-formed grammatical sentences but will instead be on the continuous ability to produce text that is appropriate to the moment.

References

