The Effect of Tutor Feedback in Language Acquisition Models

Jens Nevens
Promotor: Dr. Katrien Beuls
Advisor: Dr. Michael Spranger

Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene
jens.nevens@vub.ac.be
katrien.beuls@vub.ac.be

Abstract

This work investigates the role of tutor feedback in agent-based models of lexicon acquisition where one agent (the tutor) is teaching another agent (the learner) a particular language. An important aspect of lexicon learning is referent resolution, i.e. understanding what object in the environment is referred to by a given word. We compare two dominant paradigms in this respect: *interactive learning* and *cross-situational learning* - which differ primarily in the role of social feedback, such as pointing or gaze, in solving the referent uncertainty problem. In the former, the learner not only observes objects and words, but also receives social, pragmatic feedback. This clearly restricts the number of referents of the words. In the latter, the learner only receives objects and words. Here, the referent uncertainty problem is solved either by storing a single hypothesised referent, until evidence for the contrary is presented, or by statistical analysis of the co-occurrences of objects and words. We opt for the latter approach as it seems to align well with empirical findings.

Almost all models in lexicon learning fall into either of these two categories. However, little work has been done to compare these two paradigms systematically, except for Belpaeme and Morse (2012), who show that cross-situational learning is slower in the continuous-meaning domain due to the absence of social feedback. Importantly, real life interactions between caregiver and child probably do not fall into either of the two categories but are really combinations of full, little or no social feedback. Consequently, learning algorithms must be able to deal with situations where there is feedback interleaved with interactions where there is no feedback. In this work, we propose a new *mixed* paradigm that combines the two paradigms. This new paradigm allows to test algorithms in experiments that combine no feedback and social feedback. We control the presence or absence of social feedback in consequent interactions.

To deal with this mixed feedback setting, we extend existing learning algorithms in a new way and show how they perform with respect to the KNN approach of Belpaeme and Morse (2012) and other prototype-based approaches. Our learning algorithms use prototypes to estimate the referent shown by the tutor during subsequent training interactions. The algorithms differ in the way their

internal prototypes are updated in the presence and absence of social feedback. Our prototypes are continuous-valued and multi-dimensional and can represent various sensorimotor spaces – in our case colours. After training, we measure the communicative abilities of the learner, both in language production and language understanding, by a number of testing interactions. The algorithms are evaluated first in a purely simulated environment where the objects are randomly generated. Afterwards, a grounded environment is used. This was created in other work by humanoid robots observing scenes with physical objects. These scenes emulate a more structured and realistic learning environment. We also modify the distribution of the object's features (colours) in the grounded environment and investigate how the agents respond. In both environments, the agents are situated in contexts of multiple objects and use single words to refer to an entire object. Finally, we perform a study on a large number of parameters that all influence some aspect of language acquisition, e.g. context size, size of the tutor lexicon, total number of objects in the world and word production strategy of the tutor. We investigate how the effects of these parameters change with respect to social feedback and the proposed learning algorithms.

Our results suggest that the effect of social feedback is double. Not only does it allow the agent to learn faster, i.e. to reach the same level of success needing less interactions, the communicative abilities of the agent also become better with increasing presence of social feedback. Furthermore, we observe how the communicative success of the agents scales with respect to the social feedback in the novel mixed feedback setting. This is, however, dependent on the learning algorithm as some algorithms benefit from the presence of only some feedback more than others. Concerning the grounded world, our results paint a mixed picture. Performance in grounded worlds can be better than in simulated worlds, but this depends on the combination of algorithm and statistics of the environment. The agents are able to benefit from the additional structure present in the more realistic, grounded environments, but require at least some social feedback to do so. Finally, we deem it important not to underestimate the role of the tutor in language acquisition models and in language learning in general. In our experiments, we found that the way the tutor structures the world for the learner (i.e. the tutor's word production strategy) can have a profound impact on the learner's communicative success. This is especially true in the absence of social feedback.