

Computational Linguistics as an Applied Science

Pius ten Hacken

University of Wales Swansea, UK

p.ten-hacken@swan.ac.uk

Abstract

An applied science is opposed on the one hand to an empirical science and on the other to technology. It shares with the former its concern for explanation and with the latter its concern for solving practical, real-life problems. A prototypical example is medicine, where the purpose of research activity is both to find cures and to explain how they work. Whereas computational linguistics (CL) can also be pursued as an empirical science or used as mere technology, it is most interesting as an applied science. As an applied science CL is confronted with four basic questions about the problems to be solved: • *Problem identification*: What is the range of acceptable input and what is the type of output to be produced? • *Evaluation*: How can it be determined that the system produces the correct output and thus solves the problem? • *Problem decomposition*: How can the problem be analysed into components? • *Knowledge selection*: Which types of knowledge should be used to solve the component problems? It is important to address the questions in the order they are listed. The history of MT shows that when the question of knowledge selection is given priority over the others, there is a serious risk of ending up with a non-well-formed problem. The way explanation works in CL is largely parallel to its role in medicine. The basis of explanation is a model of communication applied to a particular situation. The object of explanation includes the wellformedness of the problem selected, its analysis in terms of the model, and the adequacy of the solution method. As a consequence, statistical methods have a mainly heuristic function.