EPISTEMIC IMPORTANCE AND THE LOGIC OF THEORY CHANGE

Peter Gärdenfors
Lund University and
Stanford University

1. INTRODUCTION

In an ideal world science would be, if not definitive, then at least cumulative. New laws and facts would be added constantly to well established theories and their consequences would be determined. However, as modern philosophy of science has so painstakingly shown us, science is full of revisions and setbacks.

The logic of an expanding theory is well-known and fairly straightforward. If a theory has been given an axiomatic basis or at least is thought of as a deductive system, then expanding the theory by a new proposition (a new law or a new fact), consists simply in adding the proposition to the theory and then assuming that all logical consequences of the new proposition fused with the old theory are also in the new theory.

On the other hand, the logic of theory revision and theory contraction is much less understood. Some philosophers of science, notably Kuhn and Feyerabend, have claimed that there is no logic in large-scale theory changes, alias paradigm shifts, because they are
intrinsically irrational processes. Here I will not be concerned with such large-scale theory changes but rather with revisions and contractions of theories within a paradigm or research program. The two most well-known accounts of the development within paradigms are Lakatos' 'research programmes' and Sneed's 'dynamics of theories'.

Lakatos (1970) conceives of a 'research programme' as producing "a series of theories $T_1, T_2, T_3, \ldots$ where each subset theory results from adding auxiliary clauses to ... the previous theory in order to accommodate some anomaly, ..." (p. 118). But accommodating an anomaly which conflicts with the older theory means that some part of the older theory has to be given up which entails that the series of theories do not represent an expanding science, but rather a series of revisions (and perhaps sometimes contractions). In general outline, Sneed's (1971) conception of 'the dynamics of theories' is similar, the main difference being that he has a more elaborate metatheory of the mathematical structure of the theories involved. It would take us too far to present Lakatos' and Sneed's metatheories in greater detail (for a lucid presentation of Sneed's theory, see Stegmüller (1976), and for a brief presentation, see Gärdenfors (1984a)). However, on this view of the dynamics of theories it is interesting to search for some rules that govern which parts of an old theory are retained when the conflicting principles of a new theory are added. Both Lakatos and Sneed separate out a 'core' of a theory which is not supposed to change during the series of revisions. Apart from this, they do not seem to say much about what determines which parts of the old theory to retain and which parts to retract. (For some schematic remarks on 'mistakes' in science, see Sneed (1971), p. 263 and pp. 285–286.)

The aim of this article is to outline a model of theory revision and theory contraction that will be a generalization of the process envisioned by Lakatos and Sneed. The key ideas are as follows: