CONDITIONALS AND COUNTERFACTUALS IN PROLOG

J.Ph. Hoepelman (IBM-IWBS,Heidelberg,FRG)

A.J.M. van Hoof (FhG-IAO,Stuttgart,FRG)

INTRODUCTION.

In our contribution for COLING 88 (HH '88), we introduced the concept of dialogical failure in the framework of dialogue games, by defining rules for the operator F, where FA is to be interpreted as "A is not winnable in this stage of the dialogue". We showed that F can be employed in the formulation of provisional implications. Provisional implications can be considered as the dialogical interpretation of defaults.It was shown that our solution works well for a range of examples. We concluded our paper with the announcement of a treatment of counterfactuals in terms of dialogical failure and of an implemented theoremprover for conditionals and counterfactuals based on our theoretical developments.In this note we will briefly describe the considerations leading to our treatment of counterfactuals and introduce the implementation of a theoremprover for conditionals as a subsystem of PROLOG. For a more detailed description of the formal properties of the system we refer to (HH.91).

I. A DOUBLE CHANGE OF LOGICAL ROLES

1. If it were the case that A,
   then it would be the case that C

is usually called a counterfactual, because its antecedent is pretended not to have the truth-value it "really" has. Tichy (T.84) shows that none of the prominent theories on counterfactuals is successfull on all members of a set of mostly very plausible testcases, and wonders why it is, that in settling the validity of a counterctual, nobody ever refers to such matters as world similarity and other well known paraphernalia of traditional counterfactual theory. We take seriously Tichy's suggestion to look for the use of counterfactuals and formulate a semantics for counterfactuals not in terms of "truth" and "falsehood", but rather in terms of what is done and not done in a dialogue in which a counterfactual appears. One thing which simply is "not done" when discussing a counterfactual is the following: Suppose you forward the thesis "if kangaroos had no tails, they would topple over" (L.73) to somebody who has already admitted that kangaroos do have tails. Suppose your adversary accepts the invitation to discuss, and takes the antecedent of the counterfactual as a temporary additional concession. Then it would be completely out of order for you to claim to have won the the discussion on the ground that your adversary has now "contradicted" himself.However, if contradicting oneself - as an opponent - in a discussion is no longer a reason for losing that discussion, then we seem to be playing according to rules which are similar to those for minimal calculus. Define negation as implication of some absurd statement, eg. a contradiction, "%":

2. $\neg A = A \text{ -> } \%$

In intuitionistic and in classical logic the proponent who utters the absurd statement loses against any thesis whatsoever. Not so in minimal calculus: here the opponent, having uttered the absurd, loses only if he, in a later move, attacks the absurd brought forward by the proponent. However, counterfactuals cannot simply be treated as implications in minimal calculus: Suppose Jones steps on the brake (B), and is alive (A), and that is all we assume or admit. We will not accept the counterfactual "If Jones would not step on the brake, then he would not be alive" as winnable (i.e. holding) under these circumstances. But in minimal calculus we have

3. $A,B \ ?_{min} \ \neg B \text{ -> } \neg A \ = \text{ yes}$

Suppose we add to the concessions A and B a concession to the effect that stepping on the brake is the only reason for Jones' being alive ($\neg B \text{ -> } \neg A$) . Now we will not want the counterfactual "If Jones would not step on the brake, then he would be alive" ($\neg B \text{ -> } A$) to be winnable. But it is, since in minimal calculus we have

4. $Q,C \ ?_{min} \ A \text{ -> } C \ = \text{ yes}$

In the usual minimal games the opponent, having admitted C has no opportunity to bring any additional reasons into play which would allow him to retract C after accepting the antecedent A. We therefore need two things: a treatment of negation which is even weaker than that of minimal calculus,and the introduction of an opportunity for the opponent to make use of his own concessions as exception rules. The second of these is easily effectuated: by introducing the fail- operator twice we cause a first change of roles which gives the opponent, now as a temporary proponent, the opportunity to bring additional concessions into play. The second fail operator then restores the initial order of roles. What we get is

5. $F(A \rightarrow F(C))$

Our counterfactual becomes,intuitively, "You,the opponent,will fail in showing that C fails, after A has been added to the concessions".

2. WEAKENING NEGATION

We obtain a system with negation which is weaker than minimal calculus negation, by assuming that there is not just one absurd statement, but possibly infinitely many. Definition 2. implicitly considers the absurd as a function taking formulae A to formulae %(A) under the assumption that %(A) = %(B) for any A and B. If we drop this assumption, which is actually a very strong one, we get a family of logics, for which the only axiom governing negation is

6. $(A \rightarrow \neg A) \rightarrow \neg A$

Valerius (V.90) calls this family "most minimal calculus" and shows that adding the assumption that %(A) = %(B) for any A and B is equivalent to

7. $(A \leftrightarrow B) \leftrightarrow (\%(A) \leftrightarrow \%(B))$

and brings us back to minimal calculus. Our final analysis of counterfactual implications "if it were the case that A, then it would be the case that B" now is

8. $F_{kin}(A \rightarrow F_{kin} B)$

where the subscript "kin" refers to the fact that the checking dialogues induced by the fail operator, are conducted according to the rules of classical games, but for the fact that negation is handled by the rules for most minimal calculus. We will demonstrate that this formalization leads to satisfactory results on all of the examples presented in (T.84).

3. IMPLEMENTATION IN PROLOG

The prover is implemented as a PROLOG subsystem. One distinguishes between the syntax of the data in the database, and the syntax of queries as in PROLOG. As for data, the program accepts facts and rules. Apart from the usual operators "," and ";", there are "neg" for negation, "< <" for provisional, non-monotonic implication, "<-" for ordinary PROLOG implication and "= >" for the counterfactual.For atomic statements the program accepts standard PROLOG syntax. They can also be built_ins, which have to be declared in order to be accessible to the meta-interpreter. The meta-interpreter is called by "success/1" and "success/2". success/1 takes a query. success/2 takes a query as first argument and a list of additional facts which can be used in the proof in addition to the facts in the database. The implementation of the meta-interpreter makes use of the PROLOG internal database facility. Interpreted date are stored in six internal databases: literal/pos, literal/neg, if/pos, if/neg, provif/pos, provif/neg.

Further code is stored in the normal PROLOG code space. In contradistiction to the PROLOG interpreter, the meta-interpreter performs loop checking. A further additional feature is a consistency check. If the goal is not a built_in,meta-interpretation searches through the internal databases in a PROLOG-like head_matching search.It determines whether it has to match a positive or a negative head. Then a database search is started in the following order: facts,monotonic rules, non-monotonic rules. Apart from the results on counterfactuals mentioned above, the prover works well on a range of cases of default reasoning, including "double diamonds", hierarchies of predicates and relevant implication.

BIBLIOGRAPHY

HH.88 Hoepelman,J.Ph.,van Hoof,A.J.M.:
The Success of Failure. The Concept
of Failure in Dialogue Logics and its
Relevance for Natural Language Semantics.
Proceedings of Coling '88.
Budapest,pp.250-255

HH.91 Hoepelman,J.Ph.,van Hoof,A.J.M.:
Two-Role,Two-Party Semantics:
Knowledge Representation,
Conditionals and Non-Monotonicity.
Oxford University Press,to appear.

L.73 Lewis,D.:
Counterfactuals, Oxford,1973

T.84 Tichy,P.:
Subjunctive Conditionals:
Two parameters vs Three.
Philosophical Studies 45,1984,pp. 147-174.

V.90 Valerius, R.:
Die Logik von Rahmen- und Stopregeln
in Lorenzen Spielen.
Diss. University of Stuttgart,1990.

end