Representing Counterfactual Conditionals

T. J. M. Bench-Capon
Department of Computer Science
University of Liverpool
England

Abstract

Counterfactual conditionals are important in the context of knowledge based systems since they play an important role in the knowledge elicitation process. This paper gives an account of counterfactuals which conforms to their use in such contexts, and which lends itself to effective implementation in knowledge based systems which use logic programming as their representational paradigm.

Key Words

Counterfactuals, Knowledge Representation, Logic Programming.

Introduction

The representation of counterfactual conditionals, by which is meant a conditional with its antecedent in the subjunctive mood and understood to be false, such as "if Napoleon had invaded England, we would speak French today", is important for the construction of knowledge based systems. If we observe someone teaching a person the kind of task for which we might use such a system, we will notice that he passes on a good deal of his expertise in the form of counterfactual conditionals. This is because not every contingency will arise in practice, and so he will need to say what different actions would have been required to meet different, not actual, circumstances. An important aspect of the knowledge being conveyed is the precise way in which circumstances alter cases. Moreover, the learner will also make heavy use of counterfactual conditionals, both to clarify what he being told, and to check his understanding. Such conditionals are, therefore, by no means a recondite topic, but one on which it is necessary to have a view, and a means of handling, when building almost any knowledge based system.

A satisfactory treatment of counterfactuals will need to achieve the following. It will need to provide a way of representing them as true so that knowledge expressed by means of them can be incorporated in the knowledge base of a system. Second it will need to provide a means of evaluating them, so that it can be checked whether or not they are consequences of a given knowledge base. Finally it is desirable that the treatment should faithfully reflect the way in which counterfactuals are used and reasoned with in practice. This paper is an attempt to achieve all three of these aims. In the paper I shall pay particular attention to their representation in logic programs, both because this is my favoured representation paradigm, and because I believe that it provides a context where where the computational treatment can be seen most easily.

It must be emphasised that there is at present no generally agreed and well understood way of representing counterfactuals. The reason that counterfactuals are not well understood may stem from one of two sources; either the problem is one of a lack of general understanding of the logic of counterfactuals in the philosophical literature, or the problem is that of providing a computational interpretation of a philosophical solution. The philosophical analysis which has proved most inspirational for workers in AI is based on a possible worlds approach to counterfactuals as exemplified by Stalnaker [4] and David Lewis [1]. Such treatments seem to us unsatisfactory, for reasons advanced elsewhere [3], and we wish to base our account on an alternative analysis of counterfactuals, which we believe is closer to the practical use of this construction.

The Philosophical Idea

First let us look at the nature of counterfactuals, and the role they play in reasoning, by considering some of the standard examples from the philosophical literature. As stated above a very popular philosophical treatment of counterfactuals is in terms of possible worlds. I shall write a counterfactual "if P were the case then Q would be the case" as P ~> Q. Broadly speaking the possible worlds approach is to say that P ~>Q is true if in the closest possible world, or set of possible worlds, in which P is the case, Q is the case.

A major deficiency of this approach is that it does not explain the ambiguities that attend counterfactuals. Thus for example

CF1 "if Caesar had been in charge in Korea he would have used the A-bomb"

seems plausible, but no more so than

CF2 "if Caesar had been in charge in Korea he would not have used the A-bomb".

In fact we could well imagine contexts in which we would be prepared to accept either of these counterfactuals as true. Any account of counterfactuals must therefore explain how this can be the case. Possible worlds analysis does so by claiming that the two cases trade on a difference in the function which selects the closest possible world(s) and says that we must determine the appropriate selection function from the context. The notion of a selection function has not, however, been made sufficiently clear for this to be a useful notion in logic programming (or for that matter in human reasoning).

Mackie [2], and certain other philosophers, have suggested that counterfactuals should not be construed as statements at all, but rather as elliptical arguments, and that the assertion of a counterfactual is the presentation of such an argument. Of course, an argument cannot be true or false it can only be valid or invalid, and so one might say that with this approach it is impossible to determine or represent the truth of a counterfactual. However, we can accept that by using a counterfactual sentence a person presents an argument without accepting that he does not at the same time make a statement. Counterfactuals, like arguments, may be mentioned or used. Where an argument is merely mentioned, or presented, there is no commitment to the truth of the premises. Where an argument is used however, its premises are asserted rather than merely entertained; thus the use of an argument does entail the making of a statement, namely the conjunction of the premises. Moreover, since we may mention an invalid argument, there is a further statement made when an argument is used sincerely, to the effect that the conclusion does indeed follow from the premises. Considered so, knowingly to use an invalid argument is not only akin to lying, it really is lying. Counterfactuals allow an argument to be presented in a particular way, that is, arguments which can be used sincerely with knowledge of the falsity of, one premise, namely the antecedent of the counterfactual, and with no explicit statement of the other premises. Thus, in using a counterfactual, one does not assert any specific premises, but does assert that there is a set of premises such that the consequent of the counterfactual follows from those premises and the antecedent, and conversationally implies the ability to make at least one set of such premises explicit if required to do so.

Of course, not any premises whatsoever will be acceptable. For although the antecedent does not need to be accepted as true since its subjective mood insulates it from such a requirement, the other premises are not protected in the same way. Thus it would seem that the other premises are required to be true. Moreover the set of premises should be required to be minimal in the sense that Q should not be provable from P + any subset of the premises.

This account gives a good explanation of the ambiguity of counterfactuals. Since there may be a number of sets of true premises which could form the basis of an appropriate argument then the counterfactual, which leaves the precise nature of the premises unspecified, could be supported by any of them, but if the assertion of the counterfactual is challenged the suppressed premises must be made explicit, thus disambiguating the counterfactual.

Also it should be noted that P ~> -Q is not the same as -(P ~> Q). This is because the first statement expresses the fact that there is a set of acceptable premises S, and P+S -> -Q, whereas the second says that there is no set of acceptable premises S such that P + S -> Q.

Now let us see how the truth of a counterfactual is considered in practice. Suppose I assert the counterfactual

CF2 If Caesar had been in charge in Korea he would not have used the A-bomb

When challenged I offer the suppressed premises

SP1 Caesar only used technology he understood

SP2 Caesar did not understand the A-bomb.

Now any one who wishes to deny the truth of CF2 as disambiguated by SP1 and SP2 would, on the above account as it stands, have either to deny SP1 or SP2 or the implication from the antecedent of CF2 and SP1 and SP2 to the consequent of CF2. But in practice there is another resource. He can deny the counterfactual by offering as a reason another counterfactual, such as,

CF3 If Caesar had been in charge in Korea he would have understood the A-bomb

offering as justification

SP3 Caesar always understood the latest technology.

Now the defender of CF2 can rebut CF3 by showing SP3 to be false, perhaps by pointing to some historical facts such as:

R1 Caesar did not understand how Greek fire worked

Of course, to maintain the defence of SP1 it would have to be the case that Caesar never used Greek fire.

Alternatively if nothing suitable to fulfill the role of R1 is at hand the defender of CF2 must rebut CF3 by using a counterfactual like

CF4 If Caesar had been in charge in Korea he would not always have understood the latest technology,

perhaps citing as suppressed premises

SP4 It requires an IQ of 168 to understand the A-bomb
and
SP5 Caesar had an IQ of 149.

Again the denier of CF2 could answer with a counterfactual such as

CF5 If Caesar had been in charge in Korea he would have had an IQ of at least 178,

giving as suppressed premises

SP6 Caesar was in the top .1% of the IQ distribution, and

SP7 the IQ distribution at the time of Korea was such that the top .1% had an IQ of least 178.

The debate could continue far beyond this, but we shall assume that it terminates here.

The points to note are:

A counterfactual can be denied in three ways
1) by denying one of the suppressed premises
2) by denying that the consequent follows from the antecedent together with the suppressed premises
3) by providing another counterfactual with the same antecedent and a denial of a suppressed premise as consequent, which does not rely on the falsity of the antecedent as a suppressed premise.

Because use of method 3 requires the production of further suppressed premises, this sort of rebuttal can itself be rebutted in any of the same three ways. This leads to a regress which can only be halted by the production of facts (not counterfactuals) which entail the denial of one of the suppressed premises.

We should also note that a suppressed premise cannot itself be a counterfactual. Thus if I wish to argue that

If Caesar were alive today he would understand logic programming,

I cannot provide the suppressed premise:

If Caesar were alive today I would explain logic programming to him.

The true counterfactual that is being expressed is

If Caesar were alive today he would understand logic programming because I would explain logic programming to him.

which is true iff

If Caesar were alive today I would explain logic programming to him and he would understand it.

The use of the counterfactual as a suppressed premise is illegitimate for the same reasons as the transitivity of counterfactuals fails, as will be discussed later, namely because the antecedent of one of the two counterfactuals may provide a counterfactual rebuttal of a premise required by the other counterfactual.

The account of counterfactuals in terms of arguments neatly explains why examples such as

CF5 If Berlioz and Verdi were compatriots Berlioz would have been Italian

CF6 If Berlioz and Verdi were compatriots Verdi would have been French

have been so puzzling in the literature. The suppressed premise in CF5 is

Verdi was Italian

Now this can be rebutted by CF6 with suppressed premise

Berlioz was French

which can in turn be rebutted by CF5! The debate is thus radically circular. Well therefore is such a thing called a paradox of counterfactual implication. If we wish to avoid this paradox we must resort to the imposition of a further restriction on the counterfactuals which can be used to rebut a counterfactual; namely that such a counterfactual would not lead to this undesirable circularity. This restriction is simply that the original counterfactual would not be, but for this restriction, itself a rebuttal of the rebutting counterfactual. It is quite in order to impose this kind of restriction, since arguments are designed to be persuasive; therefore any technique of argumentation which cannot be persuasive can be ruled out. We are not,however, forced to rule such rebuttals out of order; we may wish instead to allow them, and so allow neither CF5 nor CF6 to be true, rather than both.

At this point we should return to the question of the non-equivalence of P~>-Q and -(P~>Q). If P~>-Q then there will be a set of statements

S true in KB for which P+S->-Q. Suppose now that there is another set of statements true in KB, T, such that P+T -> Q. Now it will be the case that P+S implies the falsity of at least one member of T, thus P~>-T. Thus it will be the case that there is a counterfactual rebuttal of P~>Q (and, of course, by the same line of reasoning, of P~>-Q). Thus it is the case that P~>-Q implies -(P~>Q), but the reverse is not true, since there may simply fail to be a suitable set of premises either for P~>Q or P~>-Q.

To summarise: P ~> Q is true iff
1) there is some set of premises $S_1 \ldots S_n$ such that Q follows from P together with $S_1, \ldots, S_n$ and from no subset of this set of premises
2) the premises $S_1 \ldots S_n$ are true
3) It is not the case that P ~> -S_i for i = any 1 to n.

This, of course, means that a counterfactual cannot be definitively established as true unless a proof of the non-existence of the counterfactuals of the form of condition 3 is possible. This will not in general be the case, nor is the way in which counterfactuals are argued for in practice. In general a counterfactual will be accepted as true in a given context if no counterfactuals of the form of the third condition can be found. Thus the negation in 3 is generally accepted as negation by failure, within a circumscribed context.

This account of counterfactuals as arguments explains the three perverse properties of counterfactuals, namely the failure of contraposition, the lack of transitivity and the non-monotonicity.

Failure of Contraposition

CF8 If the power hadn't failed the dinner would have been on time

does not imply

CF9 If dinner had been late the power would have failed

since the dinner could have been late for other reasons.

CF8 represents the argument

The power was working
Nothing else made the dinner late
therefore, the dinner was on time.

Now in the case of CF9 we must assume that the dinner was on time and nothing made the dinner late. Given the thousand and one things that might make the dinner late we cannot find a good premise for an argument from the dinner was late to the power failed since the required premise, "only a power failure could have made the dinner late" is simply false. The point is that the two counterfactuals require a different state of the world for them to be counterfactuals, CF8 where the power failed and the dinner was late, CF9 where the power did not fail and the dinner was on time, so that different facts are available as suppressed premises in the different contexts of use.

Failure of transitivity

The classic argument here is that the following two counterfactuals are true

CF10 If Hoover had been born in Russia he would have been a Communist

CF11 If Hoover had been a Communist he would have been a traitor does not imply

CF12 If Hoover had been born in Russia he would have been a traitor.

Again using the argument account it is obvious that transitivity would fail

CF 10 uses a suppressed premises such as Hoover adopted the prevailing ideology and the prevailing ideology in Russia is Communism

CF11 uses the suppressed premises Hoover was the head of the FBI and anyone who is a Communist and head of the FBI is a traitor.

Clearly these premises cannot be combined as would be required by CF12 since we would have a good counterfactual rebuttal; if Hoover had been born in Russia he would not have been head of the FBI. The transitivity fails because the antecedent of one counterfactual counterfactually implies the falsity of a premise required by the other counterfactual.

Non-Monotonicity

The sort of problem here is that we may have P~>Q and P&R~>-Q.

The usual example is

CF13 If Boris had come the party would have been lively
CF14 If Boris had come and Anna had come the party would have been dreary.

Clearly it is the case that any party at which both Boris and Anna are present is dreary. Thus the absence of Anna is a suppressed premise for CF13. But this is not available in CF14, and so the argument there fails. In general if S is a suppressed premise on which P~>Q relies, then P&-S~>Q will be false.

Application to Logic Programming

Let us assume we have a logic program consisting of a set of clauses, KB. How would we evaluate the truth of P~>Q?

First we would need to attempt to find a set of premises S such that P+S->Q. This would be done by searching the data base for the set of clauses with Q as head and bodies which contain, or could be unfolded to contain, P. The clauses other than P in the bodies (unfolded as necessary) would be a set of suitable sets of premises for the counterfactual. Adopting the Closed World Assumption we would be able to say that there were no other sets of premises.

Now we would need to evaluate each member of the set of premises to determine that KB -> S.

Lastly we would need to evaluate the counterfactuals P~>-S, hoping that they would fail, or be of a kind that would be illegitimate because representing a circular argument.

An example

Let us, for an example return to the question of Caesar and the A-bomb.

Suppose KB is

```
used(abomb,X):-available(abomb,X),
               ruthless(X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
ruthless(caesar).
```

We wish to see whether

```
commandKorea(caesar) ~>
                used(abomb,caesar).
```

To prove used(abomb,caesar) we must prove available(abomb,caesar), ruthless(caesar)

This unfolds to

commandKorea(caesar), ruthless(caesar)

which expresses the conditions in terms of the antecedent.

Therefore the suppressed premise is ruthless(caesar) which follows from KB.

There is no way to prove not ruthless(caesar), so there is no counterfactual commandKorea(caesar) ~> - ruthless(caesar) thus the counterfactual is true.

But it may be that KB is not correct. While

```
used(abomb,X):-available(abomb,X),
               ruthless(X)
```

may happen to be true given those to whom the abomb was available, it may not be true absolutely. We may then wish to express KB differently:

```
used(abomb,X):-available(abomb,X),
                    ruthless(X), not careful(X).
used(abomb,X):-
        available(abomb,X), ruthless(X),
              careful(X),understood(abomb,X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
ruthless(caesar).
careful(caesar).
```

Now we het two sets of suppressed premises which unfold to

commandKorea(caesar), ruthless(caesar), not careful(caesar)

and

commandKorea(caesar), ruthless(caesar), careful(caesar), understood(abomb,caesar)

Now the third member of the first set fails so that cannot be the set of suppressed premises required for the truth of the counterfactual. Similarly the fourth member of the second set fails, so that is not a suitable set of premises. Thus there is no suitable set of premises and the counterfactual falls.

But now consider the KB

```
used(abomb,X):-available(abomb,X),
                    ruthless(X), not careful(X).
used(abomb,X):-
            available(abomb,X), ruthless(X),
              careful(X),understood(abomb,X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
ruthless(caesar).
careful(caesar).
understood(X,caesar):- available(X,caesar).
```

Now we can challenge the failure of the fourth member of the second set of premises because if we unfold the definition of understood(X,caesar) to get it in terms of the original antecedent we get the modified premise set

commandKorea(caesar), ruthless(caesar), careful(caesar), commandKorea(caesar)

and all of these succeed. This action parallels (given negation as failure as the proof of a negated proposition) the success of the counterfactual

commandKorea(caesar) ~>
                    understood(abomb,caesar).

But again there is the objection that although Caesar understood all the technology that was in fact available to him, he would have been incapable of understanding the abomb. Thus KB should be

```
used(abomb,X):-available(abomb,X),
                    ruthless(X), not careful(X).
used(abomb,X):-available(abomb,X),
                    ruthless(X),careful(X),
                    understood(abomb,X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
ruthless(caesar).
careful(caesar).
understood(X,caesar):-available(X,caesar),
             capableOfUndersanding(X,caesar).
capableOfUndersanding(abomb,X):-
                          iq(X,Q),Q>175.
iq(caesar,140).
```

Now the set of suppressed premises is

commandKorea(caesar), ruthless(caesar), careful(caesar), commandKorea(caesar), capableOfUnderstanding(abomb,caesar)

and the last of these will fail.

The reader will notice how this discussion parallels the informal argument presented above. The important difference, however, resides in the way the rebuttal by counterfactual operates. Since we use negation as failure to prove the negation of required premises, we can achieve the effect of the counterfactual rebuttal by unfolding the suppressed premises where it is possible to express them in terms of the original antecedent by so doing.

Thus the actual algorithm for deciding a counterfactual is to find a set of goals which must be provable from KB by finding clauses with the consequent as head and unfolding the body as far as is necessary to obtain any clauses in the body which can be expressed in terms of the antecedent. If no set of goals so obtained succeeds from KB then the counterfactual is false. Incidentally, of course, this use of negation as failure to establish the falsity of consequents does mean that -(P~>Q) is equivalent to P~> -Q, within the circumscribed context represented by the KB.

The observant reader will notice at this point that the above account is the same as saying that a counterfactual P~>Q is true iff KB+P -> Q. This is a somewhat surprising result, since this would represent the most naive way of evaluating counterfactuals imaginable. But there are reasons both why it is not generally applicable to unrestricted logic and why it is applicable to horn clauses augmented by negation as failure.

The reasons why it cannot be used for an unrestricted logic are as follows. Firstly since the use of the counterfactual suggests that P is false it is reasonable to think that -P is part of KB. Then it will be the case that any Q whatsoever is provable from KB + P (anything follows from a contradiction), which is simply vacuous. In general we cannot rely on KB not containing -P, and so we must remove -P from KB before we add P in order to evaluate Q. But this will not always be enough; it may be that -P, whilst not explicitly occurring in KB is provable from KB. Then again, any Q whatsoever could serve as the consequent. We ought therefore to modify KB so as to ensure the non-provability of -P from KB before attempting to prove the consequent. But even then we are not finished, for -Q may be derivable from KB. If this is the case, then we would have a proof that if the counterfactual were the case then P would not be the case, since if Q follows from P and -Q then -P. As in the case where -P is derivable from KB it makes no sense to add P to such a database. Last and worst for this line (because hardest to detect) is that some fact, R may be derivable from KB, and yet its negation be derivable from KB+P. Then again, any Q would be derivable and there would be the usual problems associated with this. In passing we may note that these difficulties in determining which modification to make to the KB, are those which attend the production of a similarity function on the possible worlds account.

But none of these problems arise if we restrict ourselves to the PROLOG subset and treat negation as failure. If P~>Q is a counterfactual and P is false, this does not imply the presence of -P in KB, but rather the absence of P and any means of proving P. Therefore we do not have to worry about removing -P or the means of proving -P from KB because the addition of P will achieve both these aims. Similarly -Q is derived from KB by failure and so if Q is derivable from KB+P -Q will not be derivable. So too with the other fact R. It is simply the case that the all the problems arose from the possibility of deriving both a statement and its negation from KB+P, whereas this is never possible if we interpret negation as failure and do not record negated statements in the database.

Thus in the case where we have only horn clauses and negation as failure we can evaluate the truth of a counterfactual by temporarily adding the antecedent (or where the antecedent is a negated statement, temporarily removing that statement) and attempting to show the consequent.

Application to the Representation Of Legislation

We can illustrate the way the above account of counterfactuals can be used by looking at an example drawn from the representation of a fragment of legislation as a logic program. Let us consider a real-life example. The Supplementary Benefits Act (1976) defines entitlement to Supplementary Benefit.

```
X entitled to Supplementary Benefit if
     X aged 16 or over
     X's resources are
           insufficient for his requirements
     X in Great Britain
     X satisfies other
                  conditions

X entitled to Supplementary Benefit if
     X aged 16 or over
     X's resources are
           insufficient for his requirements
     X abroad in certain circumstances
     X satisfies other conditions

X abroad in certain circumstances if
     not X in Great Britain
     CF(X in Great Britain, X
        entitled to Supplementary Benefit)
     X was entitled to Supplementary
              Benefit before going abroad
```

One condition a person must satisfy to be so entitled is that he or she is either in Great Britain or, if abroad, in certain specified circumstances. One of these circumstances is that the person "would, but for his absence, be entitled". There are two clauses for entitled to supplementary benefit and either of them could, in principle, form the basis of an argument presented in the counterfactual. Only one is, however, suitable for our purposes because the antecedent X in Great Britain does not occur in the second clause or any transformation of that clause achieved by repeated unfolding.
Note that it is mentioned, but not used within the counterfactual operator. Moreover, we could not use the counterfactual as a suppressed premise, because as was stated earlier no counterfactual can be used as a suppressed premise - least of all the counterfactual under consideration itself.

Thus the argument that the counterfactual presents must be the first clause. The set of suppressed premises is thus the conditions in the body of that clause other that the one that is the antecedent of the counterfactual, plus the clause itself. The truth of the counterfactual therefore depends on the truth of the suppressed premises

```
     X aged 16 or over
     X's resources are
           insufficient for his requirements
     X satisfies other conditions.
```

No counterfactual rebuttal is available since none of the goals unfold to give X in Great Britain as a goal.

In the case of a counterfactual contained within legislation it is arguable that the argument presented must itself derive from legislation. If we adopt this position it becomes more plausible to accept the closed world assumption for these cases.

Asserting Counterfactuals

Now that we have a good understanding of how we will evaluate counterfactuals, we can say what we need to do to record the truth of a counterfactual. Let us return to the Caesar example.

```
used(abomb,X):-available(abomb,X),
                  ruthless(X), not careful(X).
used(abomb,X):-available(abomb,X),
                  ruthless(X), careful(X),
                  understood(abomb,X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
ruthless(caesar).
careful(caesar).
understood(X,caesar):-available(X,caesar),
              capableOfUndersanding(X,caesar).
capableOfUndersanding(abomb,X):-
                        iq(X,Q),Q>175.
iq(caesar,140).
```

It will be remembered that we arrived at this KB as a result of modifying the starting KB which was

```
used(abomb,X):-available(abomb,X),
                  ruthless(X).
available(abomb,X) :- commandKorea(X).
commandKorea(smith).
```

ruthless(caesar).

by a number of steps which allowed or disallowed certain counterfactuals. This is just the process that we must go through if we are told that a counterfactual is true (or false). In some cases this involved the modification of rules in KB so as to place extra conditions on the satisfaction of the consequent, in other cases we expanded the definition of certain predicates. Conceivably we might also have been obliged to add extra facts. What we are doing here is either adding facts so as to ensure the satisfaction of suppressed premises, or adding (or changing) rules which can then serve as the basis for the argument which underpins the counterfactual.

Thus if we are told that a counterfactual is true we must ask whether there is an argument that would support the counterfactual. If there is not, we must provide a rule which will provide such an argument. If an argument exists, but fails due to the falsity of a suppressed premise, we must add the appropriate premise as a fact. If it fails as the result of a rebutting counterfactual, we must falsify that rebutting counterfactual, either by causing a suppressed premise of the rebutting counterfactual to be false, or by providing a counterfactual which will rebut the rebutting counterfactual. The process is illustrated by the process used in the caesar example, which in its final state established that if Caesar had been in charge in Korea he would not have used the A-bomb. If we were now told that this counterfactual is false we would have to further modify KB, perhaps by changing the value of Caesar's IQ to 180,or by giving additional rules about the development of IQ that would enable the rebutting counterfactual in the original informal argument.

Naturally, we shall not wish the modifications to KB to be arbitrary, and in the building of a real system, this might require careful knowledge elicitation to establish just why and how the counterfactual is supposed to succeed, and how and why potential rebuttals of the counterfactual fail.

Conclusion

In this paper we have offered a general analysis of counterfactuals to which their representation must conform. We have also shown that in special cases, of which a PROLOG database is one, this analysis may be satisfied in an extremely straightforward way. That this method corresponds to an (almost) absurdly naive approach to the implementation should not disappoint us; the analysis explains why it works for the special cases we wish to treat, and without the analysis its use would be dubious in the extreme. Further the analysis shows us why it would not work if applied to less restrictive theorem provers, and points the way to the kind of method that would be required to deal with counterfactuals in such context.

REFERENCES

[1] Lewis, D., *Counterfactuals.* Blackwell, Oxford, 1973.
[2] Mackie, J.L., *Counterfactuals and Causal Laws*. In R.J.Butler (ed) *Analytical Philosophy,* Blackwell, Oxford, 1962.
[3] Routen,T.W., and Bench-Capon, T.J.M., *Counterfactuals In Logic Programs.* DOC Report, Imperial College, London. 1986.
[4] Stalnaker, R., *A Theory of Conditionals*, in N.Rescher (ed), *Studies in Logical Theory,* Blackwell, Oxford, 1968.

Acknowledgement

I would like to express my thanks to Tom Routen, with whom I have had several stimulating discussions about counterfactuals and their treatment in logic programming.