A Metatheoretical Characterization of Negation as Finite Failure

Giovanni Criscuolo
(1)Dipartimento Scienze Fisiche
Mostra d'Oltremare pad.19, 80125 Napoli, Italy
e-mail: vanni@na.infn.it

Very often in science concepts and notions which for a long period of time have been considered safe and clear must be more or less radically reformulated in order to answer to pressures and exigencies arising either in the same field they originated or in related disciplines. Being science a conservative enterprise, whenever possible some of the constraints proper of the old notion are simply relaxed and a more liberal notion arises.This is exactly what happened in formal logic for the notion itself of logic. Indeed until a few years ago practically everyone would have agreed that a logic is given by:

1. One language and its set of wffs.
2. A derivability relation between sets of wffs and wffs such that the following properties hold:
3. $A \vdash A$ (reflexivity);
4. $S \vdash A$ implies $S, T \vdash A$ (monotonicity);
5. $S, B \vdash A$ and $S \vdash B$ imply $S \vdash A$ (cut).

Point 2. has been abandoned in linear logic where a derivability relation between multisets of wffs and wffs has been introduced. Nonmonotonic logics, first introduced in A.I. to cope with common sense reasoning,are now commonly studied by logicians and philosophers, thus violating point 4. Even point 1. has been recently abandoned: following a suggestion of R. Weyrauch [?], multilanguage systems have been introduced in [?]. The idea is that different theories with different languages can interact each other and export information by means of bridge rules. A bridge rule is such that if a certain formula A can be deduced in a theory T_1 then a certain other formula B can be derived in theory T_2. Due to the presence of such bridge rules the original theories T_1 and T_2 are tranformed in an effective way into two new enlarged theories T_1' and T_2'. When the language of the second theory contains an appropriate metalanguage for the first theory one obtains an OM pair. The trasformation under which undergoes an OM pair when the bridge rules are the well known reflection rules, [?] [?], or some of their more simple and immediate generalizations, have been studied in [?]. The first paragraph of this work contains the definition of OM pair while in the second some of the results proved in [?] are illustrated. Such technical results will be used in the final paragraph to give a simple characterization of negation as finite failure.

1 OM pairs

A formal metatheory is a theory about another theory called the object theory. Each proposed metatheory must have names for syntactical objects of its object theory and must contain some distinguished predicates intended to express some syntactical properties of the object theory. Of course it is up to the proposer of the metatheory to prove its correctness and/or completeness. Everyone has its own metatheory and this makes difficult to compare different approaches and to develop a systematic study of object-meta relations. In [?] the situation has beeen completely reversed. By using bridge rules as a mean to implicitely express the connection between the object theory and the metatheory one can begin a systematic study of object-meta relations.The set of bridge rules that have been studied in [?], called reflection rules, is given in fig.1.

$$[\mathrm{Rup_r}]\bullet(A)A \quad [\mathrm{Rup}]\bullet(A)A \quad [\mathrm{Rup^n_r}]\neg\bullet(A)\neg A \quad [\mathrm{Rup^n}]\neg\bullet(A)\neg A$$

$$[\mathrm{Rdw_r}]A\bullet(A) \quad [\mathrm{Rdw}]A\bullet(A) \quad [\mathrm{Rdw^n_r}]\neg A\neg\bullet(A) \quad [\mathrm{Rdw^n}]\neg A\neg\bullet(A)$$

RESTRICTIONS: rules labeled with index r are applicable if and only if the premiss does not depend on any assumptions in the same theory.

Fig. 1. Reflection rules

We now give the formal definition of OM pair

Definition 1 *Given a logical language L, its pure propositional metalanguage $\bullet(L)$ is the language whose set of atomic wffs is the following:*

$$\{\bullet(A) : A \textit{is an } L\textit{-wff}\}$$

Definition 2 *An OM pair is given by* $O, M, (\mathrm{RR})$ *where:*

- $O = L, \Omega_O, \Delta_O$ *is an object theory in the language* L *with axioms* Ω_O *and set of deduction rules* Δ_O.
- $M = \bullet(L), \Omega_M, \Delta_M$ *is a metatheory in the pure propositional metalanguage of* L *with axioms* Ω_M *and deduction rules* Δ_M.
- (RR) *is any subset of reflection rules of figure* **??**.

In the following L will be a propositional language and Δ_O and Δ_M are the classical natural deduction rules.

The definition of derivation in an OM pair is straigthforward. One can see such a derivation as made up by parts, each of which is either an objective subderivation or a subderivation in the metalanguage appropiately interconnected by reflection rules.

$\mathrm{TH}(O)$ ($\mathrm{TH}(M)$) is the set of all classical consequences of the objective axioms (of the meta-axioms), while $\mathrm{TH_{OM}}(O)$ ($\mathrm{TH_{OM}}(M)$) is the set of all objective formulae (metaformulae) provable in the OM pair. Of course $\mathrm{TH}(O) SS \mathrm{TH_{OM}}(O)$ and $\mathrm{TH}(M) SS \mathrm{TH_{OM}}(M)$. Here is a proof of $\bullet(AB) \bullet (A) \bullet (B)$ in any OM pair containing $\mathrm{Rup_r}$ and Rdw in its set of reflection rules.

$[]\bullet(AB) \bullet (A) \bullet (B)[]\bullet(A) \bullet (B)[\mathrm{Rup_r}]\bullet(A)[]A[\mathrm{Rdw}]AB\bullet(AB)[\mathrm{Rup_r}]\bullet(B)[]B[\mathrm{Rdw}]AB\bullet(AB)$

2 Some results on OM pairs

Our investigation developed along three main lines:

1. Comparing the strenght of different sets of reflection rules.
2. For a given set of reflection rules trying to characterize $\mathrm{TH_{OM}}(O)$ and $\mathrm{TH_{OM}}(M)$ in terms of O and M.
3. For a given set of reflection rules trying to understand what relation in the object theory does $\bullet(A)$ expresses.

About point 1, I will only mention that a partial order on sets of reflection rules can be obtained using the following definition:

Definition 3 $(\mathrm{RR})_1 \leq (\mathrm{RR})_2$ *if, for any* O *and* M, $S \vdash A$ *in* $O, M, (\mathrm{RR})_1$ *implies* $S \vdash A$ *in* $O, M, (\mathrm{RR})_2$.

The strongest combination turns out to be $\mathrm{Rup} + \mathrm{Rdw} + \mathrm{Rup_r^n}$, all the other rules being derivable from these.

Turning to point 2 above, we have that $\mathrm{TH_{OM}}(O)$ can be expressed in terms of $\mathrm{TH_{OM}}(M)$. The problem therefore becomes: can we axiomatize $\mathrm{TH_{OM}}(M)$?

If S is a set of objective wffs let:

$$\bullet(S) = \{\bullet(A) : A \in S\} \tag{1}$$

Notice that, no matter what O, M and (RR) are we have that $\Omega_M SS\mathrm{TH}_{\mathrm{OM}}(M)$ and that, if (RR) cotains Rup_r, then $\bullet(\mathrm{TH}(O))SS\mathrm{TH}_{\mathrm{OM}}(M)$.

The problem therefore becomes the following: if $P_1, \ldots, P_n$ are parameters and (RR) is a subset of reflection rules containing Rup_r, to find a $\bullet(L)$ schema $\Phi[P_1 \ldots, P_n]$ such that, for any object theory O and metatheory M, one has

$$\mathrm{TH}_{\mathrm{OM}}(M) = \mathrm{TH}(\Omega_M \cup \bullet(\mathrm{TH}(O)) \cup \{\Phi[A_1, \ldots, A_n] : A_i \in L\}) \tag{2}$$

Notice that if this is possible for any set (RR) of reflection rules than the approach via bridge rules would be equivalent to the usual approach of axiomatizing the metatheory. Surprising enough this is not the case. Indeed in [**?**] we proved that the answer is negative in case (RR) = $\mathrm{Rup}_\mathrm{r} + \mathrm{Rdw}_\mathrm{r}$, by giving a counterexample. This is by no way a trivial result if one realizes that the restricted reflection rules above are the ones most commonly used in the literature. When Rdw_r is relaxed to Rdw then usually an axiomatization can be given. We refer the interested reader to [**?**].

Point 3 of course is the more interesting one. Given any couple O,M and a fixed (RR), one would like to express $\mathrm{TH}_{\mathrm{OM}}(O)$ in terms of the models of $\mathrm{TH}_{\mathrm{OM}}(M)$. If one can further specify what are the models of $\mathrm{TH}_{\mathrm{OM}}(M)$, then the meaning of $\bullet(A)$ can be understood. This can be easily done when one realizes that there is a one to one correspondence between sets of objective formulae and models of $\bullet(L)$. Indeed given a model m of $\bullet(L)$ let

$$\| \bullet \|_m = \{A \mid \ m \models \bullet(A)\}$$

Viceversa a given set of formulae S implicitely defines the model that evaluates $\bullet(A)$ to true iff A is in S. By an extension of O we mean any set of formulae closed under modus ponens that contains $\mathrm{TH}(O)$. A set of objective sentences is maximal if it contains A or $\neg A$ for any formula A. In [**?**] the following theorem has been proved:

Theorem 4 *Let* OM *be an OM pair composed of O and M connected by the set of reflection rules* (RR)*;*

1. *If* (RR) *is* $\mathrm{Rup}_\mathrm{r} + \mathrm{Rdw}$*, then m is a model of* $\mathrm{TH}_{\mathrm{OM}}(M)$ *if and only if* $m \models \Omega_M$ *and there exists a set of wffs* ΓSSL *such that* $\| \bullet \|_m = \mathrm{TH}(O + \Gamma)$*;*

2. *If* (RR) *is* $\mathrm{Rup}_\mathrm{r} + \mathrm{Rdw} + \mathrm{Rup}_\mathrm{r}^\mathrm{n}$*, then m is a model of* $\mathrm{TH}_{\mathrm{OM}}(M)$ *if and only if* $m \models \Omega_M$ *and there exists a set of wffs* ΓSSL *such that* $\| \bullet \|_m = \mathrm{TH}(O + \Gamma)$ *and* $\mathrm{TH}(O + \Gamma)$ *is* consistent*;*

3. *If* (RR) *is* Rup+Rdw*, then m is a model of* $\mathrm{TH}_{\mathrm{OM}}(M)$ *if and only if* $m \models \Omega_M$ *and there exists set of wffs* ΓSSL *such that* $\| \bullet \|_m = \mathrm{TH}(O+\Gamma)$ *and* $\mathrm{TH}(O+\Gamma)$ *is* maximal*;*

4. *If* (RR) *is* Rup + Rdw + $\mathrm{Rup}_{\mathrm{r}}^{\mathrm{n}}$*, then* m *is a model of* $\mathrm{TH}_{\mathrm{OM}}(M)$ *if and only if* $m \models \Omega_M$ *and there exists a set of wffs* ΓSSL *such that* $\| \bullet \|_m = \mathrm{TH}(O+\Gamma)$ *and* $\mathrm{TH}(O+\Gamma)$ *is* maximal *and* consistent*;*

5. *If* (RR) *is* $\mathrm{Rup}_{\mathrm{r}}^{\mathrm{n}}$ + Rdw*, then* m *is a model of* $\mathrm{TH}_{\mathrm{OM}}(M)$ *if and only if* $m \models \Omega_M$ *and, if* $\mathrm{TH}(O)$ *is inconsistent then* $\| \bullet \|_m = \emptyset$*, otherwise* $O + \| \bullet \|_m$ *is* consistent*.*

Theorem 5 *Let* OM *be an OM pair composed of* O *and* M *connected by a set of reflection rules containing* $\mathrm{Rdw}_{\mathrm{r}}$ *and* $\mathrm{Rup}_{\mathrm{r}}$*, then* $\mathrm{TH}_{\mathrm{OM}}(O) = \bigcap_{m \models \mathrm{TH}_{\mathrm{OM}}(M)} \| \bullet \|_m$.

Therefore one can say that the models of $\mathrm{TH}_{\mathrm{OM}}(M)$ are in case 1 all the extensions of O, in case 2 all the consistent extensions of O, in case 3 all the maximal extensions of O, etc., that satisfy the axioms.

3 A characterization of negation as failure

Let us consider a logic program P where negation is implemented under the finite failure rule, i.e. $notA$ succeeds iff the prooof of A finitely fails. The finite failure of A is only a contingent property of the proof procedure strongly determined by the particular context in question, the program P. It is not invariant under extensions of the original program P and therefore it is non monotonic in nature. However for any program P the following metaproperty is certainly true:

$$\text{if } notA \text{ can be proved then } A \text{ cannot be proved} \tag{3}$$

Notice that the converse does not hold. In the program $P = A \leftarrow A$ neither A nor $notA$ can be proved. Only interpreting $notA$ as failure to prove does the converse hold. Let us consider a propositional language L' whose set of wffs are obtained by the usual logical connectives plus the unary connective not. We have now the possibility to formalize property (**??**) as a meta axiom and to study what are the constraints that such an axiom imposes on the unary connective not. Let $O, M,$ (RR) be such that O is any objective theory in the language L' while M has as its only axiom:

$$\bullet(notA) \neg \bullet (A) \tag{4}$$

and (RR) = $\mathrm{Rup}_{\mathrm{r}}$ + Rdw Now:

$$\bullet(notA) \neg \bullet (A) \equiv \neg(\bullet(notA) \bullet (A))$$

which implies, see the example in section 1,

$$\neg \bullet (notAA)$$

Consider the following derivation:

$$[] \neg \bullet (\neg(notA \neg A))[] \bot [\mathrm{Rup}_{\mathrm{r}}] \bullet (notAA) = notAA [\mathrm{Rdw}] \neg(notA \neg A) \bullet (\neg(notA \neg A)) \neg \bullet (notAA)$$

Using the theorem of the previous section, this proof tells us that for any objective theory O, no extension of O can contain $\neg(notA\neg A)$. In particular any complete extension, any truth set, must contain $notA\neg A$.In other words we can use this formula as an axiom. Notice that nothing changes if we intepret $\neg$ as intuitionistic negation. No intuitionistic theory can contains $\neg(notA\neg A)$, otherwise there would be a classical theory containing it. Intuitionistic negation satisfies $\neg A(AB)$ and $(A\neg A)\neg A$.

¿From $notA\neg A$ one derives

$$notA(AB) \tag{5}$$

but

$$(AnotA)notA \tag{6}$$

does not hold. An equivalent system based on axiom (**??**) can be found in Gabbay [**?**]. It is easy to show that,

$$\neg A \equiv AnotA$$

If (**??**) holds, than the two negations are equivalent. If one interprets $notA$ as failure to prove A then instead of axiom (**??**) one should use the meta axiom

$$\bullet(notA) \equiv \neg \bullet (A)$$

It can be easily shown that in this case (**??**) holds and the connective *not* becomes intuitionistic negation. Many people think that negation as finite failure is Nelson strong negation. Now strong negation, indicated with $-$, also satisfies the property

$$-A\neg A$$

but moreover it also satisfies:

$$\begin{gathered} --A \equiv A \\ -(AB) \equiv -A-B \\ -(AB) \equiv -A-B \\ -(AB) \equiv A-B. \end{gathered}$$

These further constraints are not a consequence of our axiom (**??**). Only the right to left implication does hold when $-$ is substituted with *not*. Although a reasonable possibility, to impose the other side of the implication is problematic. Indeed how *not* should behave strongly depends from the modalities of the computation, for example if it is serial or parallel, etc. The only thing that one can say is that finite failure, if interpreted as falsity, is a form of negation that lies in between two extremes: strong negation and intuitionistic negation.

References

[1] K. Bowen and R. Kowalski. Amalgamating language and meta-language in logic programming. In S. Tarlund, editor, *Logic Programming*, pages 153–173, New York, 1982. Academic Press.

[2] G. Criscuolo, F. Giunchiglia, and L. Serafini. A Foundation of Metalogical Reasoning: OM pairs (Propositional Case). Technical Report 9403-02, IRST, Trento, Italy, 1994.

[3] D. M. Gabbay. *Semantical Investigations in Heytings Intuitionistic Logic*. D. Reidel, 1981.

[4] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics (or: how we can do without modal logics). *Artificial Intelligence*, 65:29–70, 1994. Also IRST-Technical Report 9110-07, IRST, Trento, Italy.

[5] R. Weyhrauch. Prolegomena to a Theory of Mechanized Formal Reasoning. *Artificial Intelligence*, 13(1):133–176, 1980.