Nonmonotonic reasoning and modal logic, from negation as failure to default logic

Philippe Balbiani

NONMONOTONIC REASONING AND MODAL LOGIC,

FROM NEGATION AS FAILURE TO DEFAULT LOGIC

Philippe Balbiani
Institut de Recherche en Informatique de Toulouse, université Paul Sabatier,
118 route de Narbonne, F-31062 Toulouse Cedex

abstract : We present a modal characterization of two well-known nonmonotonic formalisms : the negation as failure rule and default logic. The semantics of logic programming with the negation as failure rule is described through the definition of a modal completion. In modal logic K4, this completion characterizes provability in logic programming with respect to SLDNF-resolution while in modal logic Pr (the modal logic of provability) it characterizes unprovability in logic programs.

keywords : nonmonotonic reasoning, logic programming, modal logic

1. PRESENTATION

Negation as failure is a non-monotonic rule of inference. It introduces provability and unprovability features inside a linear resolution-based deduction mechanism. It makes reference to a logic program inside the program itself. See Clark (1978), Jaffar, Lassez and Lloyd (1983) and Lloyd (1984) for deeper investigations in logic programming with the negation as failure rule.

Default logic uses inference rules like A : B / C which allows us to add C to our current knowledge database whenever A belongs to that database and B is consistent with that database. Since something which is consistent with a set of belief is not always consistent with a superset of that set of belief, default logic is a nonmonotonic logic. See Reiter (1980) and Besnard (1989) for deeper investigations in default logic.

The purpose of this report is to study some properties of negation as failure and default logic. It gives a modal translation of logic programs. It shows this translation is sound and complete with respect to the negation as failure rule. It gives a modal translation of default theories and shows the soundness of this translation with respect to the provability in default logic. It studies the relationship between logic programming with negation as failure and default reasoning. It defines two very similar linear resolution-based mechanisms. The first mechanism can be proved to be equivalent to SLDNF-resolution - see Lloyd (1984) for a detailed study of SLDNF-resolution. The second mechanism enlightens us on the notion of provability in a fragment of default logic. It characterizes from a proof theory point of view the relationship between logic programming with negation as failure and default logic.

Though we will only consider ground logic programs, the results of sections 3 and 4 can be extended to the predicate case.

2. MODAL LOGIC

We will use the modal logics $K4^{\pm}$, Pr^{-} and $Pr^{\pm}$. Their language is based on a finite set VAR of propositional variables or atoms - atoms and negation of atoms will be called literals. The language includes the operators $\neg$, $\vee$, $\wedge$ and $\rightarrow$, the modal operators [+] and [-], and the rule : if F is a well-formed formula (wff) then so are [+]F and [-]F.

Modal logic $K4^{\pm}$ possesses the propositional calculus axioms and rules of inference plus the axioms : $[s](A\rightarrow B)\wedge[s]A\rightarrow[s]B$ (s=+,-) and $[s]A\rightarrow[s'][s]A$ (s=+,- and s'=+,-) and the rule : if F is a theorem then so are [+]F and [-]F. Modal logic Pr^{-} possesses the axioms and rules of $K4^{\pm}$ plus the axiom : $[-]([-]A\rightarrow A)\rightarrow[-]A$. Modal logic $Pr^{\pm}$ possesses the axioms and rules of Pr^{-} plus the axiom : $[+]([+]A\rightarrow A)\rightarrow[+]A$. Modal logics Pr^{-} and $Pr^{\pm}$ are variations of the modal logic of provability Pr - see Boolos (1979) or Smorynski (1984) - which is famous because of its relationship with the godelian concept of provability in arithmetic.

The semantics of these modal logics is defined in term of Kripke models. A Kripke model is composed of :

(1) a non-empty set W of possible worlds,
(2) two binary relations R^{+} and R^{-} defined over the members of W and
(3) an application v which associates to any pair (w,A) -w being an element of W and A being a well-formed formula- an element of {0, 1} such that :

(3)(1) $v(w,\neg A)=1-v(w,A)$,
(3)(2) $v(w,A\wedge B)=v(w,A)\times v(w,B)$ and, for any s in {+,-},
(3)(3) $v(w,[s]A)=\min\{ v(w',A) : wR^{s}w' \}$.

A wff A is said to be valid in a model (W,R^+,R^-,v) when $v(w,A)=1$ for every element w of W.

The adequation results - see Balbiani (1991) for detailed proofs - between the axiomatic and the semantics of the modal logics $K4^{\pm}$, Pr^- and $Pr^{\pm}$ are as follows :

Theorem : The well-formed formula A
(1) is a theorem of $K4^{\pm}$ iff it is valid in any model (W,R^+,R^-,v) where R^+ and R^- are mutually transitive, that is to say : $wR^{s'}w''$ whenever $wR^{s}w'$ and $w'R^{s'}w''$,
(2) is a theorem of Pr^- iff it is valid in any model (W,R^+,R^-,v) where W is finite, R^+ and R^- are mutually transitive and R^- is irreflexive, that is to say : there is no element w of W such that wR^-w,
(3) is a theorem of $Pr^{\pm}$ iff it is valid in any model (W,R^+,R^-,v) where W is finite, R^+ and R^- are mutually transitive and R^+ and R^- are irreflexive.

Following the filtration method detailed in Hughes and Cresswell (1984), it can be showed that modal logics $K4^{\pm}$, Pr^- and $Pr^{\pm}$ possess the finite model property, that is to say :

Theorem : For any formula A of $K4^{\pm}$ (respectively : Pr^- and $Pr^{\pm}$) such that $\neg A$ is not a theorem of $K4^{\pm}$ (respectively : Pr^- and $Pr^{\pm}$), there is a model (W,R^+,R^-,v) of $K4^{\pm}$ (respectively : Pr^- and $Pr^{\pm}$) such that W is a finite set of possible worlds and, for some w in W, $v(w,A)=1$.

Consequently, and for obvious reasons that will not be much more detailed, $K4^{\pm}$, Pr^- and $Pr^{\pm}$ are decidable.

3. LOGIC PROGRAMMING

3.1. Preliminary definitions

A goal is a set of literals. A clause is a formula like $l_1 \wedge l_2 \wedge \ldots \wedge l_m \rightarrow A$ where A is an atom and each l_i is a literal. A logic program is a set of clauses. For every program P and for every atom A, P(A) will denote the set $\{L : L \rightarrow A \in P\}$ and $P(\neg A)$ will denote the set $\{ L : (\forall l \in L)(\exists L' \in P(A))(-l \in L') \} \cap \{ L : (\forall L' \in P(A))(\exists l \in L)(-l \in L') \}$, where, for every atom B, $-B=\neg B$ et $-\neg B=B$.

Let N be an application which associates to (P,l) - P being a program and l a literal - the program $P'=N(P,l)$.

An N-derivation from the goal L and the program P will be any sequence (L_0,P_0), (L_1,P_1), ..., (L_n,P_n) such that :

$$n\geq 0,\ L_0=L,\ P_0=P \text{ and, for every i ranging from 0 to n-1,}$$
$$(\exists l\in L_i)(\exists L'\in P(l))(L_{i+1}=(L_i\setminus\{l\})\cup L')(P_{i+1}=N(P_i,l)).$$

An N-refutation of L in P will be an N-derivation (L_0,P_0), (L_1,P_1), ..., (L_n,P_n) from L and P such that $L_n=\emptyset$.

N_1 will denote the application which associates P to (P,l). N_1-refutations are of course the refutations that would be obtained in an idealized version of PROLOG. N_2 will denote the application which associates P to (P,A) and $P\setminus A$ to $(P,\neg A)$, where $P\setminus A=P\setminus\{L\rightarrow A:L\in P(A)\}$. If $N=N_1$ then, using Fitting (1985) and Kunen (1987), it can be proved that :

Theorem : L is N_1-refutable in P if and only if (iff) there is an SLDNF-refutation of $P\cup\{L\rightarrow\bot\}$, that is to say iff L succeeds in P with respect to a refutation-based mechanism using the negation as failure rule.

3.2. Translation in modal logic

We translate "L is N_1-refutable in P" into the formula $P^*\rightarrow\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B$, [+] and [-] being the modal operators of $K4^{\pm}$; P^* being a translation of P in $K4^{\pm}$. In the modal logic $K4^{\pm}$ we have previously defined, $P^*= P_1\wedge P_2$ where

$P_1=\wedge_{A\in VAR}\,[+](\vee_{L\in P(A)}(\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B)\rightarrow A)$,
$P_2=\wedge_{A\in VAR}\,[-](\wedge_{L\in P(A)}(\vee_{B\in L}[-]B\vee\vee_{\neg B\in L}[+]B)\rightarrow A)$

is both sound and complete with respect to N_1-refutation, that is to say :

Theorem : L is N_1-refutable in P iff $P^*\rightarrow\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B$ is a theorem of $K4^{\pm}$.

Proof : The proof of the if-part is done by induction on the length of the N_1-refutation of L in P. The proof of the only if-part is done by induction on the depth of the closed tableau of $\{\neg(P^*\rightarrow\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B)\}$. See Balbiani (1991) for further details.

Examples :
The modal translation of $P=\{\neg A\rightarrow B\}$ is the formula
$P^*=[-]A\wedge[+]([-]A\rightarrow B)\wedge[-]([+]A\rightarrow B)\wedge[-]C$.
$\{\neg A,B,\neg C\}$ is N_1-refutable in P and $P^*\rightarrow[-]A\wedge[+]B\wedge[-]C$ is a theorem of $K4^{\pm}$.

The modal translation of $P=\{\neg B\rightarrow B,\neg A\rightarrow B,\neg C\rightarrow A\}$ is the formula $P^*=[+]([-]C\rightarrow A)\wedge[-]([+]C\rightarrow A)\wedge[+]([-]B\vee[-]A\rightarrow B)\wedge[-]([+]B\wedge[+]A\rightarrow B)\wedge[-]C$.
$\{A,\neg C\}$ is N_1-refutable in P and $\{B\}$ et $\{\neg B\}$ are not N_1-refutable in P and $P^*\rightarrow[+]A\wedge[-]C$ is a theorem of $K4^{\pm}$ and $P^*\rightarrow[+]B$ and $P^*\rightarrow[-]B$ are not theorems of $K4^{\pm}$.

As for non-provability in logic programs with negation, in $Pr^{\pm}$ P^* is both sound and complete, that is to say :

Theorem : L is not N_1-refutable in P iff $P^*\rightarrow\vee_{B\in L}[-]B\vee\vee_{\neg B\in L}[+]B$ is a theorem of $Pr^{\pm}$.

Note : The previous property is equivalent to :

L is N_1-refutable in P iff $P^*\wedge\neg(\vee_{B\in L}[-]B\vee\vee_{\neg B\in L}[+]B)$ is $Pr^{\pm}$-satisfiable

Proof : The if-part is proved by induction on the length of the N_1-refutation of L in P. The only if-part is proved by induction on the depth of the Kripke model of $Pr^{\pm}$ which satisfies $P^*\wedge\neg(\vee_{B\in L}[-]B\vee\vee_{\neg B\in L}[+]B)$. See Balbiani (1991) for further details.

These results characterize the notions of provability and unprovability in logic programs with the negation as failure. They use the fact that the formula $[+]A\vee[-]A$ - which logically translates the alternative between the refutability of A and the refutability of $\neg A$ - is not a tautology of modal logic. It proves - especially for the notion of unprovability - that modal logic is able to represent the deduction characteristics in logic programs with negation. The last result can be explain by the fact that for any goal N_1-refutable in some program there exists a finite derivation tree which possesses an empty leaf and by the fact that Kripke's models of $Pr^{\pm}$ are finite irreflexive trees. This is the point which has led us to translate default theories into the modal logic Pr^{-}.

4. DEFAULT LOGIC

4.1. Preliminary definitions

A default theory T is composed with a set of Horn clauses Ax(T) and a set D(T) of rules (or defaults rules) like $A_1\wedge A_2\wedge...\wedge A_m : \neg B_1, \neg B_2, ..., \neg B_n / C$ where $m\geq 0$, $n>0$ and where$A_1, A_2, ..., A_m, B_1, B_2, ..., B_n$ and C are atoms.

An extension of T will be any fixpoint of Γ_T where for every set E of wff, $\Gamma_T(E)$ is the smallest set E' of wff such that $Ax(T)\subseteq E'$, $Th(E')=E'$ - $Th(E')$ being the deductive closure of E' - and : $(\forall\ F{:}\neg B_1\neg B_2...\neg B_n/C \in D(T))(\text{if } (F\in E')(\forall i=1..n)(B_i\notin E) \text{ then } C\in E')$.

Examples : The theory ({A}, {A : ¬B / C, true : ¬C / B}) possesses two extensions:Th({A, B}) and Th({A, C}). The theory ({A}, {true : ¬B / C, C : ¬A / B, C: ¬B / D}) possesses only one extension:Th({A, C, D}). The theory ({}, {true : ¬B / B}) possesses no extension.

4.2. Translation in modal logic

T(A) will denote the set $\{L{:}L\to A\in Ax(T)\} \cup \{\{A_1, A_2, ..., A_m\}\cup\{\neg B_1, \neg B_2, ..., \neg B_n\} : (A_1\wedge A_2\wedge...\wedge A_m{:}\neg B_1\ \neg B_2\ ...\ \neg B_n /A \in D(T))\}$. Let T be a theory which possesses almost one extension.

The formula T^* of Pr^- where

$T^*= T_1\wedge T_2$,
$T_1=\wedge_{A\in VAR}\,[+](\vee_{L\in T(A)}(\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B) \to A)$ and
$T_2=\wedge_{A\in VAR}\,[-](\wedge_{L\in T(A)}(\vee_{B\in L}[-]B\vee\vee_{\neg B\in L}[+]B) \to A)$

is sound for provability in T. That is to say :

Theorem : For every goal L, if there is an extension E of T such that $(L\cap VAR\subseteq E)(\forall B)(\neg B\in L\to B\notin E)$ then $T^*\to\wedge_{B\in L}[+]B\wedge\wedge_{\neg B\in L}[-]B$ is a theorem of Pr^-.

Let T=({A}, {A : ¬B / C, true : ¬C / B}). T possesses two extensions:Th({A, B}) and Th({A, C}). We have $T^*=[+]A\wedge[+]([+]A\wedge[-]B \to C)\wedge[+]([-]C \to B)\wedge...$
$...\wedge[-]([-]A\vee[+]B \to C)\wedge[-]([+]C \to B)$ and $T^*\to[+]A, T^*\to[+]B, T^*\to[+]C, T^*\to[-]B$ and $T^*\to[-]C$ are theorems of Pr^-. Our translation is not complete however and the previous theory T shows it well. The modal wff $T^*\to[+]B\wedge[+]C$ is a theorem of Pr^- but B and C does not belong to the same extensions. We believe however that this difficulty could be solved by another kind of translation of default theory in Pr^-.

Let T=({A}, {true : ¬B / C, C : ¬A / B, C : ¬B / D}). T possesses only one extension : Th({A, C, D}). We have :

$$\begin{aligned}T^*= &[+]A\wedge[+]([-]B \to C)\wedge[+]([+]C\wedge[-]A \to B)\wedge\\ &[+]([+]C\wedge[-]B \to D)\wedge[-]([+]B \to C)\wedge\\ &[-]([-]C\vee[+]A \to B)\wedge [-]([-]C\vee[+]B \to D)\end{aligned}$$

and $T^*\to[+]A\wedge[-]B\wedge[+]C\wedge[+]D$ is a theorem of Pr^-.

The modal logic Pr^- expresses the notion of provability in a fragment of default logic. Its main axiom $[-]([-]A\rightarrow A)\rightarrow[-]A$ expresses the uselessness of an axiom like $A\wedge L\rightarrow A$ or a default rule like $A\wedge L{:}\neg B_1,\neg B_2, \ldots,\neg B_n/A$ in a default theory.

In the case where T possesses no extension, the soundness is not assured. Let $T=(\{B\}, \{true : \neg A / A\})$. T possesses no extension. Its modal translation is $T^*=[+]B\wedge[+]([-]A\rightarrow A)\wedge[-]([+]A\rightarrow A)$. In Pr^-, we have $[-]([-]A \rightarrow [+][-]A)$. Thus $T^*\rightarrow[-]([-]A \rightarrow A)$ and $T^*\rightarrow[-]A$, though no extension of T contains A. Our translation is not able to represent the fact that the presence of the default $true{:}\neg A/A$ in a theory can make this theory inconsistant in the sense that it possesses no extension.

5. LOGIC PROGRAMMING AND DEFAULTS

The previous result does not give to our fragment of default logic a complete formalization in modal logic. Now we show an interesting result concerning the relationship between logic programming with negation and default reasoning. Provability in default logic is in fact equivalent to the notion of provability based on N_2-refutation. Semantical relations between the autoepistemic logic of Moore (1985) and stratified logic programs - see Apt, Blair and Walker (1988) for precise investigations inside this particular class of program - has been presented by Gelfond (1987). Bidoit and Froidevaux (1988) has given similar results for default logic. Lifschitz (1988) has studied the relations between stratified programs and circumscription. Our approach will do no restriction concerning the stratifiability of logic programs.

Let T be a theory of default logic. Let $P(T)=\{ L\rightarrow A : (A\in VAR)(L\in T(A)) \}$. An N_2-refutation $(L_0,P_0), (L_1,P_1), \ldots, (L_n,P_n)$ of L in P(T) is said to be total when $\{ A : (A\in VAR)(P(T)(A)\neq\emptyset) \}$ is included in $\{A:\{A, \neg A\}\cap(L_0\cup L_1\cup \ldots\cup L_n)\neq\emptyset\}$.

The following result establishes the equivalence between the notion of provability in default logic and N_2-refutability in P(T).

Theorem : If there is a total N_2-refutation of L in P(T) then there is an extension E of T such that $(L\cap VAR\subseteq E)(\forall B)(\neg B\in L\rightarrow B\notin E)$ and if there is an extension E of T such that $(L\cap VAR\subseteq E)(\forall B)(\neg B\in L\rightarrow B\notin E)$ then there is an N_2-refutation of L in P(T).

Examples : The theory $T=(\{\}, \{true : \neg A / B, true : \neg B / A\})$ possesses two extensions : $Th(\{A\})$ and $Th(\{B\})$. The associated program is $P(T)=\{\neg A\rightarrow B, \neg B\rightarrow A\}$. The sequence $(\{\neg A\}, P(T)), (\{B\}, P(T)\backslash A), (\{\neg A\}, P(T)\backslash A), (\{\}, P(T)\backslash A)$ is a total N_2-

refutation of $\{\neg A\}$ in P(T). As for the goal $\{A, B\}$ - which is not provable in T - it does not possess any N_2-refutation in P(T).
The theory $T=(\{B\}, \{true : \neg A / A\})$ possesses no extension. The associated program is $P(T)=\{\emptyset \rightarrow B, \neg A \rightarrow A\}$. The sequence $(\{B\}, P(T)), (\{\}, P(T))$ is an N_2-refutation of $\{B\}$ in P(T) which is not total, and that corresponds to the fact that $\{B\}$ is not provable in T.

6. PERSPECTIVES AND REFERENCES

We have presented a modal semantic for negation in logic programming. This semantic used a very particular modal logic, namely Pr, which is related to the notion of provability in arithmetic. We believe that non-monotonic formalisms and this modal logic share common specificities and that some of the properties of Pr are those the main non-monotonic logic are searching for. For full first order logic programs with negation, previous results have been extended, see Balbiani (1991).

REFERENCES

Apt K. R., Blair H. A. and Walker A. 1988, Towards a theory of declarative knowledge, in Foundations of Deductive Databases and Logic Programming (J. Minker, Ed.), Morgan Kaufmann Publishers, Los Altos, CA, 89-148.

Balbiani P. (1991), Une caractérisation modale de la sémantique des programmes logiques avec négation, thèse de l'université Paul Sabatier, janvier 1991.

Balbiani P. (1991), A modal semantics for the negation as failure and the closed world assumption rules, to be presented at the eighth Symposium on Theoretical Aspects of Computer Science, Hamburg, february 1991.

Besnard P. 1989, An introduction to default logic, Springer-Verlag.

Bidoit N. and Froidevaux C. 1988, More on stratified default theories, Proc. of the European Conf. on Artificial Intelligence, 492-494.

Boolos G. 1979, The unprovability of consistency, Cambridge University Press.

Chellas B. F. 1980, Modal logic : an introduction, Cambridge University Press.

Clark K. L. 1978, Negation as failure, in Logic and Databases (H.Gallaire and J.Minker, Eds) Plenum Press, New York, 293-322.

Fitting M. R. 1985, A Kripke-Kleene semantics for logic programs, J. of Logic Programming, 2:295-312.

Gelfond M. 1987, On stratified autoepistemic theories, Proc. of the Sixth Nat. Conf. on Artificial Intelligence (AAAI-87).

Hughes G. E. and Cresswell M. J. 1984, A companion to modal logic, London, Methuen.

Jaffar J., Lassez J.-L. and Lloyd J. W. 1983, Completeness of the negation as failure Rule, Proc. of the Eighth Int. Joint Conf. on Artificial Intelligence, Karlsruhe, 500-506.

Kripke S. 1963, Semantical analysis of modal logic, Zeitschrift für Mathematische Logik und Grundlagen der Mathematik, 9:67-96.

Kunen K. 1987, Negation in logic programming, J. of Logic Programming, 4:289-308.

Lifschitz V. 1988, On the declarative semantics of logic programs with negation, in Foundations of Deductive Databases and Logic Programming, (J. Minker, Ed.), 177-192.

Lloyd J. W. Foundations of logic programming, Springer Verlag, 1984.

Moore R. C. 1985, Semantical Considerations on Nonmonotonic Logic, Artificial Intelligence, 25:75-94.

Reiter R. 1980, A logic for default reasoning, Artificial Intelligence, 13:81-132.

Smorynski C. 1984 Modal logic and self reference, Handbook of Philosophical Logic (D. Gabbay et F. Guenthner, Eds), Reidel, 441-496.