Relating The TMS to Autoepistemic Logic

Yasushi Fujiwara
Shinichi Honiden

Systems & Software Engineering Laboratory
TOSHIBA Corporation
70 Yanagicho
Kawasaki 210, JAPAN

Abstract

Truth maintenance systems have been studied by many authors and have become powerful tools in AI reasoning systems. From the viewpoint of commonsense reasoning, Doyle's TMS seems most interesting, for it allows nonmonotonic justifications. Its semantics, however, has remained unclear. In this paper, we shall give its declarative description in terms of autoepistemic logic, a kind of nonmonotonic logic. That is, we shall exhibit a one-to-one correspondence between states acceptable to the TMS and stable expansions of autoepistemic formulas attached to justifications. Thus, the TMS turns out to be a theorem prover of autoepistemic logic. For the practical interest, our result also suggests the possibility of implementing better TMS algorithms by using the theorem proving method of autoepistemic logic.

1. Introduction

In this paper, we shall give the declarative semantics of Doyle's TMS [6]. Our method is based on autoepistemic logic defined by Moore [14]. The main result is that there exists a natural one-to-one correspondence between states acceptable to the TMS and stable expansions of the set. of autoepistemic formulas attached to the justifi cations.

In Doyle's first paper on the TMS, his main intention seemed to be to put nonmonotonic reasoning into practical use. But his description of the TMS was algorithmic and without semantics. To provide the semantics is important not only for a theoretical interest but also for a practical one in improving TMS implementations.

McDerinott and Doyle [13] attempted to give a logical background of the TMS. However, their "non-monotonic logic" has several disadvantages. Some attempts to resolve these faults have been made [12,14]. Among others, Moore's autoepistemic logic has clear semantics. In addition, it is shown that autoepistemic logic has remarkable relations to the modal logic S5 [15] and Reiter 's default logic [10]. But these studies of nonmonotonic reasoning seem to have little influence on the work on truth maintenance.

De Kleer [3] presented the ATMS architecture, a variant of Doyle's TMS. It aims at efficient search and can process multiple contexts simultaneously. The ATMS, however, can treat monotonic justifications only. There are proposals of the ATMS architecture which allow nonmonotonic justifications [4,5,7], but they also lack the semantics.

In this paper, we shall make the semantics of Doyle's TMS clear by using the technique of nonmonotonic reasoning. *The TMS is a theorem prover of autoepistemic logic.* Our result suggests the possibility of implementing better TMS and extended ATMS algorithms.

2. Autoepistemic logic

Moore [14] defined autoepistemic logic as a formal framework of beliefs of the ideally rational agent-reasoning about her own beliefs. Moore [15] further obtained alternative semantics, which is based on Krlpke semantics of the modal logic S5. In this section, we shall give a brief account of his theory.

2.1. The formalism of autoepistemic logic

The language of autoepistemic logic is that of propositional logic augmented by a unary connective L. The symbol L is intended to mean "is believed". We suppose that atomic propositions are drawn from a finite set P.

Let T be a set of autoepistemic formulas. For any propositional truth assignment V, we define the autoepistemic interpretation V? to be the truth assignment which extends V by the condition

Vi'(Lp) = 1 <=> ptT.

If an autoepistemic interpretation Vf satisfies the condition that V-p(p) -1 for all p t T, we shall call Vp an autoepistemic model of T.

We define the notions of soundness and completeness relative to this semantics.

Definition 1. Let A be a finite set of autoepistemic formulas. A set of autoepistemic formulas T is *sound* with respect to a set of premises A if and only if every autoepistemic interpretation of T in which every formula of A is true is also an autoepistemic model of T.

Definition 2. A set of autoepistemic formulas T is *seimmtically complete* if and only if every autoepistemic formula which is true in every autoepistemic model of T lies in T.

The set of beliefs that a rational agent might hold, given a set of premises A, would be semantically complete theory that is sound with respect to A.

Definition 3. A set of autoepistemic formulas T is called *a stable expansion* of a set of premises A if and only if T satisfies the following conditions:

1) T contains A.
2) T is sound with respect to A.
3) T is semantically complete.

2.2. The modal logic S5 [9,11]

The logic S5 is a kind of modal logic of knowledge. Its Kripke semantics is very simple. An S5 Kripke model is just a set of truth assignments. These truth assignments can be considered as the worlds which are possible.

Syntax. The language of S5 is syntactically identical to that of autoepistemic logic. The symbol ~L~ is often abbreviated as M.

Semantics. An S5 Kripke structure is a set of propositional truth assignments. An S5 model is a pair (V,K) consisting of a propositional truth assignment V and an S5 Kripke structure K such that V€K. The interpretation of an S5 formula p with respect to this model is given by the usual truth recursion augmented by conditions:

1) (V,K) |= p <=> V |= p if p is a propositional formula.
2) (V,K)| = Lp <=> (W,K)| = p for all WK.

We shall denote by Vk(p) the truth value of a formula p with respect to an S5 model (V,K).

2.3. Autoepistemic logic and S5 semantics

The formulation of autoepistemic logic given above is nonconstructive and makes it difficult to seek stable expansions. In [15], Moore characterized semantically complete autoepistemic theories in terms of S5 semantics. This characterization enables us to demonstrate the existence of stable expansions of given set of premises.

Theorem 1. (Moore) The following conditions are equivalent.

1) The autoepistemic theory T is semantically complete.
2) The autoepistemic theory T is given by
   T = { p | (V,K) | - p for all V€K},
for some *Sb* Kripke structure K.

By the above theorem, we get simple characterization of stable expansions, which we shall use repeatedly.

Proposition 1. Let A be a set of autoepistemic formulas. Then a stable expansion T of a set of premises A corresponds bijectively to an S5 Kripke structure K which satisfies the following conditions:

1) T - { p | (V,K) |= p for all V€K}.
2) For any propositional truth assignment V,
   VeK <=> V_T(p) = 1 for all p€A.

Proof. Let T be a stable expansion of a set of premises A and define an S5 Kripke structure K by

K = { V | VT is an autoepistemic model of T }.

From Moore's proof of the above theorem, we have

T = (p | (V,K) |= p for all V€K}.

On the other hand, as T contains A and is sound with respect to A, VT is an autoepistemic model of T if and only if V_T(p) = 1 for all p€A. Thus, K satisfies 1) and 2).

Suppose that an S5 Kripke structure K satisfies 1) and 2). By 2), if V € K, we have V_T(p) =1 for all p€A. Remark that, for V€K, V_T(q) = V_K (q) for any autoepistemic formula q. Then we have VK(P) = 1 for any peA and VcK. By the definition of T, we see that TDA. By 2), if a propositional truth assignment V satisfies the condition that VT(P) =1 for all peA, we get VeK, hence VT is an autoepistemic model of T. Thus, T is sound with respect to A. This shows that T is a stable expansion of a set of premises A. Q.E.D.

Combining Proposition 1 and the decidability of the modal logic S5 [9,11], we know that autoepistemic logic is also decidable, i.e. we can obtain all the stable expansions of the given set of premises.

3. The TMS

A TMS is a part, of the reasoning system. The reasoning system consists of a problem solver and a TMS. The problem solver transmits every inferences made to the TMS. The TMS manages justifications and answers what, data are believed ("in") and disbelieved ("out") when asked.

In this section, after giving a formulation of the TMS, we shall define a translation rule from the TMS theory to autoepistemic logic and show its properties.

3.1. The formalism of the TMS

Here we give a formulation of Doyle's TMS. For simplicity, we shall only deal with SL-justi fications.

Definition 4. *A TMS* is a triple D= (N,J,C) such that

1) N is a finite set (The elements of N will be called *nodes.*) .
2) J is a subset of $Nx2^Nx2^N$, where 2^N denotes the power set of N (The elements of J will be called *justifications.)* .
3) C is a subset of 2^N (The elements of C will be called *nogoods.)* .

Let j- (n,Ni,N2) J be a justification. Then the node n is called the consequent node of j. The subset Ni (resp. N2) of N is called the inlist (resp. outlist) of the justification j. The justification j is said to be nonmonotonic if the outlist N2 is nonempty.

The following definition of admissible states is intended to formulate states acceptable to a TMS without "circularity-check".

Definition 5. Let D= (N,J,C) be a TMS and S be a subset of N. We shall say that S is a^ *admissible state* of D if S satisfies the following conditions:

1) Let j = (n,Ni,N2) € J be a justification with Ni c S and N2cN\S. Then n lies in S.

2) Conversely, for any neS, there exists a justification j - (n,Ni,N2) eJ with Ni c S and Np. c N\S.

3) For any nogood c, c4S.

Let S be an admissible state of D. By the above definition, for any node n€S, there exists a justification j- (n,Ni,N2) such that Ni c S and N2CN\S. We shall call such j a supporting justification of n.

Doyle tried to get rid of "in" nodes supported by circular arguments. His TMS singles out one supporting justification to each "in" node and tries to ensure that the set of supporting justifications is without circularity. The following definition of well-founded admissible states is intended to formulate states acceptable to the TMS with "circularity-check". Although Doyle classified three patterns of circular arguments, we shall consider the first one only. The second type of circularity is an example of multiple-extension problem which are common in nonmonotonic reasoning. The third type of circularity is unsatisfiable one and it has no admissible state in our sense.

Definition 6. Let D and S be as in the above definition. We shall say that S is a *minimal admissible state* of D if S is an admissible state of D and any proper subset of S is not an admissible state of D.

Definition 7. Let D and S be as in the above definition. An admissible state S of D is said to be *a well-founded admissible state* ii_r for any nεs, there exists a supporting justification j_n of n such that S is a minimal admissible state of the TMS (N, { j_n I ri€ S },0) (Such a set { j_n I n *: S } is called a set of well-founded supporting justifications.).

The following example shows our definition precludes admissible states which are based on circu1 a r ar gument s.

Example 1. Let *J* - { (p, {p}, 0), (q,0, {p}) } • Then there are two admissible states, i.e. {p} and {q}, which are both minimal. The supporting justification (p,{p},0) of the node p is circular, hence {p} is not well-founded. On the other hand, the supporting justification of q is (q,0,{p}) and {q} is a minimal admissible state of {(q, 0, {p})}. Thus, the state {q} is the unique well-founded admissible state of J.

Remark. In the first version of this paper, we formulated the states acceptable to the TMS with circularity-check to be minimal admissible states. But as we can see in the above example, it is insufficient. The authors also considered yet another truth maintenance system [8] based on stratified logic programming technique [16], which accepts only the state {q} of the above example. Thus, we felt a need to reformulate states acceptable to the TMS. The definition given here is a natural formalization of Doyle's original one.

A TMS may have no well-founded admissible states and may have more than one. As long as we allow the full use of nonmonotonic justifications, there seems to be little hope to resolve this difficulty. The set of justifications of the following example has two well-founded admissible states, but there seems to be no reasonable way to select one of them as canonical. Later we shall discuss the class of the set of justifications w>ich has one and only one well-founded admissible state.

Example 2. Consider the set of justifications

J = { (p,0, {q}) , (q,0, {p}) } .

Then there are two admissible states of J, i.e. {p} and {q}. It is easily checked that both are well-founded.

The following proposition shows that well-foundedness subsumes minimality.

Proposition 2. Let D and S be as in the above definition. If S is a well-founded admissible state of D, then it is also a minimal admissible state of D.

Proof. Suppose that S is not a minimal admissible state of D. Then there exists a proper subset To of S which is an admissible state of D.

From this assumption, we shall construct an admissible state T of the TMS (N, { j_n | neS},0). Denote by Jo ~ (jn I ^{n} *=^{s} } the set of well-founded supporting justifications. Let J^ be the set consisting of justifications jn~ (n,Ni,N2) € Jo with Ni c To and Tj be the union of TQ and the set of consequent nodes of J\

T] - Tou{ n | j_n - (n,Ni,N_2) t j_0 with Ni c T_0 } .

Let J;> be the set consisting of justifications j_n = (n,Ni,N2) "' J() with N] <_ Ti and T2 be the union of To and the set. of consequent nodes of J'2

T2=Tou{ 11 I i n - (n,N],N2) <= J'0 with N_2 C T] }.

As J'2 contains J], *Ty* contains Ti . By repeating this argument, we get a chain of sets

To c Ti c.cTj c ....

Denote by T the union of all the sets Ti (i 2 0). Then it is easily checked that T is an admissible state < f the TMS (N,Jo,0).

On the other hand, as To is an admissible state of D, we see that the set of consequent nodes of J"l is a subset of To. Hence we get TQ = T] = T2=... = T. Thus, we obtain an admissible state To of the TMS (N,Jo,0), which is a proper subset of S. This contradicts the assumption that S is a well-founded admissible state of D. Q.E.D.

3.2. The TMS and autoepistemic logic

If I is a finite set of formulas, we shall denote by AI (resp. VI) the conjunction (resp. disjunction) of the formulas in I.

Let D = (N,J,C) be a TMS. By considering N as a set of atomic propositions, we associate to any justification j = (n,Ni,N2) an autoepistemic formula F(j) defined by

$$\wedge\{ Lk \mid k \in N_1 \} \wedge \wedge\{ \sim Ll \mid l \in N_2 \} \supset n.$$

Put $F(D) = \{ F(j) \mid j \in J \}$. For any subset S of N, we define an S5 Kripke structure k(S) by

$$k(S) = \{ V \mid V(s) = 1 \text{ for all } s \in S \}$$

and an autoepistemic theory t(S) by

$$t(S) = \{ p \mid (V,k(S)) \models p \text{ for all } V \in k(S) \}.$$

Remark. The above translation rule has a natural interpretation. The node is "in" if and only if the corresponding atomic proposition lies in the autoepistemic theory (the set of "beliefs"). It is interesting that the above rule is very similar to the rule used in Konolige's proof [10] of the "equivalence" of default logic [19] and autoepistemic logic.

We can now state our main result.

Theorem 2. The mapping S -> t(S) gives a bijection between the following sets.

1) The set of admissible states of D.
2) The set of stable expansions T of F(D) with Ac ¢ T for all C € C.

Proof. For any autoepistemic theory T, define a subset s (T) of N by s (T) = (nЄNInЄT).

It suffices to show the theorem for the case that C - 0. We shall proceed in three steps:

Step 1. If 5 is an admissible state of D, t(S) is a stable expansion of F(D).

Step 2. If T is a stable expansion of F(D), s(T) is an admissible state of D.

Step 3. s(t(5)) = S for any subset S of N and t(s(T)) = T for any stable expansion T of F(D).

Proof of Step 1. Let S be an admissible state of D. By Proposition 1, to prove our claim, it suffices to show that

$$Vek(S) \iff V_{t(S)}(F(j)) = 1 \text{ for all } j \in J.$$

We first show the implication =>. Let V be an element of k(S). By the fact that S is an admissible state of D, we have Vt(S) (F(j)) = 1.

For the other implication, suppose that a propositional truth assignment V satisfies the RHS. By the definition of k(S), it suffices to show that V(s) =1 for all s Є S. Let s be any element of S. As S is an admissible state of D, there exists a supporting justification j of s. Thus, $V_{t}(S)(F(j)) = V_{t}(S)(s) = V(s) = 1$. This shows the other implication.

Proof of Step 2. Let T be a stable expansion of F(D) and K be the corresponding S5 Kripke structure as in Theorem 1. Let us show that s(T) satisfies 1) and 2) of Definition 5. To show 1), let j - (n,N1,N2) be an element of J such that Nics(T) and N2 c N\s(T) . Since T is a stable expansion of F(D), any V of K satisfies the condition VT(F(J)) = 1 for all j € J. Thus, we have V(n) =-1 for V€K. This implies that n € T.

Now we show that s(T) satisfies 2) of Definition 5. Let s be an element of s(T). Suppose that there exists no justification j = (s,Ni,N2) with Nics(T) and N2CN\s(T). Let V be any element of K and W be a propositional truth assignment such that W(n) = V(n) (n # s) and W(s) = 0. Then it is easily observed that Wx(F(j)) =-1 for all j€ J. As T is a stable expansion of F(D), we have W€K by Proposition 1. Since s belongs to the set

$$S(T) - \{ p \in N \mid V(p) - 1 \text{ for all } V \in K \},$$

we have W(s) - 1, a contradiction.

Proof of Step 3. The equality s (t (S)) - S is trivial. Let T be a stable expansion of F(D) and K be the corresponding S5 Kripke structure. To prove the equality t(s(T)) - T, it suffices to show that k(s(T)) = K. Note that

$$k(s(T)) - \{ V \mid V(s) - 1 \text{ for all } s \in s(T) \}$$
$$= \{ V \mid V(s) = 1 \text{ for all node } s \in T \}.$$

The inclusion k(s(T)) is clear. Suppose that Vek(s(T)). By using the fact that s(T) is an admissible state of D, we easily see that $V_{T}(F(j))$ - 1 for all jЄJ. As T is a stable expansion of F(D), we have VЄK by Proposition 1. This shows the other inclusion. Q.E.D.

In the original definition of autoepistemic logic, all the st.abie expansions are considered to be of the same rank. For example, let A ={LpDp}. Then there are two stable expansions of A, one has p as a belief and one has not. But in the former, p is believed on tenuous bases. The proposition p is derived because Lp is derived. This is a circular argument. The latter seems more appropriate for the belief set of the rational agent. This problem is similar to the circularity-check problem of the TMS.

Konolige [10] introduced strongly grounded stable expansions to get rid of stable expansions based on circular arguments. By using his idea, we can extend the correspondence given above to the TMS with circularity-check.

To introduce the notion of strongly grounded stable expansions, we first give the definition of minimal stable expansions.

Definition 8. Let A be a finite set of autoepistemic formulas. A set of autoepistemic formulas T is called *a minimal stable expansion* of A if T is a stable expansion of A and the corresponding *5b* Kripke structure is maximal among all the S5 Kripke structures corresponding to stable expansions of A.

Then the following characterization of minimal admissible states is a direct consequence of Theorem 2.

Theorem 3. The mapping S -> t(S) gives a bijection between the following sets.
1) The set of minimal admissible states of D.
2) The set of minimal stable expansions T of F(D) with $\wedge c \notin T$ for all $c \in C$.

Proof. It is sufficient to show that, for a minimal admissible state S of D, the stable expansion t(S) of F(D) is minimal. Let K denote the S5 Kripke structure which corresponds to t(S). Suppose that t(S) is not minimal. Then there exists a stable expansion T' such that the corresponding S5 Kripke structure K^1 includes K properly. Since $\wedge c \notin t(S)$, we immediately have $\wedge c \notin T'$ Thus, T' also corresponds to an admissible state of D, a contradiction. Q.E.D.

An autoepistemic formula is said to be normal if it is of the form

$$L\alpha_1 \wedge \ldots \wedge L\alpha_m \wedge \sim L\beta_1 \wedge \ldots \wedge \sim L\beta_n \supset \gamma \quad (i, j \geq 0),$$

where α_i, β_j and y are prepositional formulas. The autoepistemic formula F(j) attached to the justification j is normal. Konolige [10] defined the notion of strongly grounded stable expansions for the set of normal autoepistemic formulas. Notice that the strongly grounded stable expansions of the given set of premises is dependent on the presentation of the formulas.

Definition 9. Let A be a finite set of normal autoepistemic formulas and T be a stable expansion of A. The set T is said to be a *strongly grounded stable expansion* of A if T is a minimal stable expansion of A', where A' is the subset of A defined as follows:

$$\{(L\alpha_1 \wedge \ldots \wedge L\alpha_m \wedge \sim L\beta_1 \wedge \ldots \wedge \sim L\beta_n \supset \gamma) \in A \mid \gamma \in T, \{\beta_i\ (1 \leq i \leq n)\} \cap T = \emptyset\}$$

The following theorem shows that states acceptable to the TMS (with "circularity-check") are completely characterized as strongly grounded stable expansions of the normal autoepistemic formulas attached to justifications.

Theorem 4. The mapping S -> t(S) gives a bijection between the following sets.
1) The set of well-founded admissible states of D.
2) The set of strongly grounded stable expansions T of F(D) with $\wedge c \notin T$ for all $c \in C$.

Proof. Suppose that S is a well-founded admissible state of D. By the definition, there exists a set of well-founded supporting justifications $\{ j_n \mid n \in S \}$. Denote by Jo the subset of J such that

$$J_0 = \{ j = (n, N_1, N_2) \in J \mid n \in S,\ N_2 \subset N \backslash S \}.$$

Remark that $j_n \in J_0$ for all n€S and that S is also a well-founded admissible state of (N,Jo,C). To show that the stable expansion t(S) of F(D) is strongly grounded, it is sufficient to show that t(S) is a minimal stable expansion of F(J0). By Proposition 2 and the above remark, we see that S is a minimal admissible state of (N, J0,C). By Theorem 3, we conclude that t(S) is a minimal stable expansion of F(J0).

Let T be a strongly grounded stable expansion of F(D). Then T is a minimal stable expansion of F(J'), where J' is the set of justifications given by $J' = \{ j = (n, N_1, N_2) \in J \mid n \in T,\ N_2 \cap T = \emptyset \}$. Thus, s(T) = { n€N I ri€T } is a minimal admissible state of J' by Theorem 3. Let J1 be the set of justifications j= (n,N1,N2) € J' with N1 = ø and S1 be the set of consequent nodes of J1

$$S_1 = \{ n \mid j = (n, \emptyset, N_2) \in J' \}.$$

Let J2 be a set of justifications j= (n,N1,,N2) $\in$ J' with N1 ⊂S1 and S2 be the set of consequent nodes of J2

$$S2 = \{ n \mid j = (n,N1,N2) \in J' \text{ with } N1 \subset S1\}.$$

As J2 contains J1, S2 contains *S1*. By repeating this argument, we obtain a chain of sets

$$S_1 \subset S_2 \subset \ldots \subset S_i \subset \ldots \subset s(T).$$

Denote by S the union of all Si (i>0). Then it is easily checked that the subset S of s(T) is an admissible state of J'. By the minimality assumption of s(T), we get S = s(T). Let n be an element of s(T) and i be the least integer such that neSj. We define the supporting justification j_n of the node n to be the justification j- (n,N1,N2) € J' such that $^{N1} \subset$si-1 (In the case that i = 0, we let S-1=ø.). It is easy to see that the set { j_n I n€s } is a set of well-founded supporting justifications of S. Q.E.D.

4. Stratified case

In general, a TMS may have no well-founded admissible state and may have more than one. But if we impose suitable restrictions on the use of outlists, we can guarantee that the TMS has a unique well-founded admissible state. In the rest of this paper, we shall discuss such a class.

The problem of nonmonotonic justifications in truth maintenance has a strong similarity to that of negation in logic programming. In the field of logic programming, there are also many attempts to extend logic programming incorporating the full use of negation [2,18]. However, all of them have severai difficulties. Especially, any positive use of negation in the presence of recursion has not been obtained [20].

Apt, Blair, Walker and Van Gelder [1,21] introduced a class of sets of clauses which prohibit recursion "through negation". From the semantic viewpoint, in such a set of clauses, we only negate propositions whose meanings are fixed beforehand. Then a "canonical" model is assigned to such a set of clauses.

We here adopt their idea. We shall consider the set of justifications which is "stratified". Our result says that a stratified set of justifications has a unique well-founded admissible state. We suppose that this result is of practical importance in the use of the TMS.

Definition 10. A partition

$J = J_1 \cup \ldots \cup J_n$ (disjoint union)

is called a *stratification* of J if the following two conditions hold:

1) If a node n occurs in a justification in *Ji,* its definition is contained within $\cup_{k \le i} J_k$.

2) If a node n occurs in the outlist of a justification in Ji, its definition is contained within $\cup_{k<i} J_k$. (The *definition* of n is the subset of J consisting of all the justifications whose consequent nodes are n.)

Then J is said to be stratified by $J_1 \cup \ldots \cup J_n$ and each Ji is called a stratum of J.

Example 3. Let $J1 = \{(p, \emptyset, \{q\}), (q, \emptyset, \{p\})\}$. Then there exists no stratification of *J1*.

Let $J_2 = \{(p, \{p\}, \emptyset), (q, \emptyset, \{p\})\}$. Then the partition $J2 = \{(p, \{p\}, \emptyset)\} \cup \{(q, \emptyset, \{p\})\}$ is the stratification of J2.

Let D= (N, J,C) be a TMS. We shall say that D is stratified if J is stratified. Our result is

Theorem 5. Let $D = (N, J, \emptyset)$ be a stratified TMS. Then D has one and only one well-founded admissible state.

Proof. Let the partition $J = J_1 \cup \ldots \cup J_n$ be the stratification of J. We first construct a well-founded state S by using this stratification.

Put So = 0. For an integer i such that 1< i< n, we shall define sets of nodes Si. Suppose that Si-1 is already defined. Define the set of justifications Ji $^{(1)}$ by

$$J_i^{(1)} = \{ j = (n, N_1, N_2) \in J_i \mid N_1 \subset S_{i-1},\ N_2 \cap S_{i-1} = \emptyset \}$$

and define the set of nodes $S_i^{(1)}$ by

$$S_i^{(1)} = S_{i-1} \cup \{ n \mid j = (n, N_1, N_2) \in J_i^{(1)} \}.$$

Define the set of justifications $J_i^{(2)}$ by

$$J_i^{(2)} = \{ j = (n, N_1, N_2) \in J_i \mid N_1 \subset S_i^{(1)},\ N_2 \cap S_{i-1} = \emptyset \}$$

and define the set of nodes $S_i^{(2)}$ by

$$S_i^{(2)} = S_{i-1} \cup \{ n \mid j = (n, N_1, N_2) \in J_i^{(2)} \}.$$

As $J_i^{(2)}$ contains $J_i^{(1)}$, $S_i^{(2)}$ contains $S_i^{(1)}$. Define $J_i^{(3)}$ by

$$J_i^{(3)} = \{ j = (n, N_1, N_2) \in J_i \mid N_1 \subset S_i^{(2)},\ N_2 \cap S_{i-1} = \emptyset \}$$

and define the set of nodes $S_i^{(3)}$ by

$$S_i^{(3)} = S_{i-1} \cup \{ n \mid j = (n, N_1, N_2) \in J_i^{(3)} \}.$$

By repeating this construction, we obtain a chain of sets of nodes

$$S_i^{(1)} \subset S_i^{(2)} \subset \ldots \subset S_i^{(m)} \subset \ldots.$$

Denote by S_i the union of all $S_i^{(m)}$ $(m \ge 0)$. Thus, we obtain S_i for all $1 \le i \le n$. We put $S = S_n$.

We now prove that S constructed above is a well-founded admissible state of D. First we show that S is admissible. Let j = (n,N1,N2) be a justification such that $N1 \subset S$ and $N2 \cap S = \emptyset$. Let us show that $n \in S$. Suppose that $j \in J_i$. Then we have $Ni \subset Si$, hence $Ni \subset S_i^{(k)}$ for some positive integer k. Because $Si\text{-}1 \subset S$, we see that j belongs to the set $Ji^{(k+1)}$, hence $n \in Si^{<k+1>} \subset Si$.

On the other hand, suppose that $n \in S$. Then there exists an integer i such that $j = (n,N1,N2) \in Ji$ with $Ni \subset Si$ and $N_2 \cap S_{i-1} = \emptyset$. By the construction of S, we see that $N_2 \cap S = \emptyset$. Thus, j is a supporting justification of s.

We now prove that S is well-founded. We associate a node $n \in S$ with its supporting justification j_n as follows: Let i be the least integer such that $n \in Si$. Then there exists an integer k such that $n \in S_i^{(k)} \setminus S_i^{(k-1)}$. By the construction, there exists a justification $j_n = (n,N1,N2)$ such that $Ni \subset Si^{(k-1)}$ and $N_2 \cap S_{i-1} = \emptyset$. It is easy to check that $\{j_n \mid n \in S\}$ is the set of well-founded supporting justifications of S.

Conversely, let T be a well-founded admissible state of D. Denote by N* the set of nodes whose definitions are contained within kii^{J}k- To show that S and T coincide, it is sufficient to prove that $S_i = S \cap N^i = T \cap N^i$ for all i by using induction on i. Notice that, by the stratifiability of J, Si and $T \cap N^1$ are well-founded admissible states of the TMS $(N^1, \cup_{j_C} iiJ\}_c, 0)$ for all i.

For i = 0, both sides are empty. Let us assume that $S_i = T \cap N^i$. We first show that $S_{i+1} \subset T \cap N^{i+1}$. It suffices to prove that $S_{i+1}^{(k)} \subset T$ for all k by induction on k. Let $n \in S_{i+1}^{(1)}$. By the construction of S, there exists a justification $j = (n, N_1, N_2)$ such that $N_1 \subset S_i$ and $N_2 \cap S_i = \emptyset$. By the assumption, we have $N_1 \subset T$ and $N_2 \cap T = \emptyset$. Thus, $n \in T$. Suppose that $S_{i+1}^{(k)} \subset T$ for an integer k. Let $n \in S_{i+1}^{(k+1)}$. Then there exists a justification $j = (n, N_1, N_2)$ with $N_1 \subset S_{i+1}^{(k)}$ and $N_2 \cap S_i = \emptyset$. By the assumption, we have $N_1 \subset T$ and $N_2 \cap T = \emptyset$, hence $n \in T$. Thus, we get $S_{i+1}^{(k+1)} \subset T$. Hence we know that $S_{i+1} \subset T \cap N^{i+1}$. But S_{i+1} and $T \cap N^{i+1}$ are both well-founded admissible states of $(N^{i+1}, \cup_{k \le i+1} J_k, \emptyset)$, hence minimal admissible states of $(N^{i+1}, \cup_{k \le i+1} J_k, \emptyset)$ by Proposition 2. Thus, we obtain $S_{i+1} = T \cap N^{i+1}$. This establishes our induction step. Q.E.D.

Remark. Theorem 5 suggests a relationship between the semantics of the TMS and that of logic programming. In fact, the fixpoint semantics can serve as yet another semantics of the TMS. In [8], the authors clarified the semantics of the basic ATMS in terms of propositional Horn logic. The correctness of the label update algorithm follows from the least fixpoint semantics. We also proposed an extended ATMS architecture based on the iterated fixpoint semantics of stratified logic programming. We hope to discuss the relationships among these results in the subsequent paper.

5. Comparison with related work

Reinfrank and Dressier [17] independently have established the relationship between the TMS and autoepistemic logic. Their result and ours are essentially equivalent. The differences between them lie mainly in technical subtleties: Our method is based on Moore's possible world formulation and theirs is based on Konolige [10]. Our definition of states acceptable to the TMS with circularity-check seems to be nearer to Doyle's original one than theirs, but they turn out to be the same.

We suppose that the uniqueness of the well-founded admissible state of the stratified TMS is suggestive to proposals of extended ATMS architectures.

Conclusion

To make nonmonotonic inference a practical technique, we must fully understand the semantics of Doyle's TMS. In this paper, we have shown that its semantics is completely described in terms of autoepistemic logic. States acceptable to a TMS correspond bijectively to stable expansions of a set of autoepistemic formulas attached to justifications. The implementation of the new TMS algorithm based on autoepistemic logic is to be explored.

References

1. Apt, K. R., Blair, H and Walker, A., Towards a theory of declarative knowledge, in Minker, J., editor, Foundations of deductive database and logic programming, pp.89-148, Morgan Kaufmann, Los Altos, 1987.
2. Clark, K. L., Negation as failure, in Gallaire, H and Minker, J., editors, Logic and data base, pp.293-322, Plenum, New York, 1978.
3. de Kleer, J., An assumption-based truth maintenance system, Artificial Intelligence 28, pp.127-162, 1986.
4. de Kleer, J., Extending the ATMS, Artificial Intelligence 28, pp.163-196, 1986.
5. de Kleer, J., A general labeling algorithm for assumption-based truth maintenance, in Proc. of AAAI-88, 1, pp.188-192.
6. Doyle, J., A truth maintenance system, Artificial Intelligence 12, pp.231-272, 1979.
7. Dressier, 0., An extended basic ATMS, Siemens Report INF 2 ARM-3-87.
8. Fujiwara, Y., Mizushima, Y and Honiden, S., On logical foundations of the ATMS, to appear.
9. Hughes, G. E and Cresswell, M. J., An introduction to modal logic, Methuen, London, 1968.
10. Konolige, K., On the relation between default theories and autoepistemic logic, Artificial Intelligence 35, pp.343-382, 1988.
11. Kripke, S., Semantical considerations of modal logic, Zeitschrift fur Mathematlsche Logic und Grundlangen der Mathematik 9, pp.67-96, 1963.
12. McDermott, D., Nonmonotonic logic II: nonmonotonic modal theories, J. ACM 29 (1) pp.33-57, 1982.
13. McDermott, D and Doyle, J., Non-monotonic logic I, Artificial Intelligence 13, pp.41-72, 1980.
14. Moore, R. C, Semantical considerations on nonmonotonic logic, in Proc. of 8-th IJCAI, pp.272-279.
15. Moore, R. C, Possible-world semantics for autoepistemic logic, in Proc. of 1984 Nonmonotonic Reasoning Workshop, pp.344-354, American Association for Artificial Intelligence, New Paltz, 1984.
16. Przymusinski, T., On the relation between logic programming and nonmonotonic reasoning, in Proc. of AAAI-88, 2, pp.444-448.
17. Reinfrank, M and Dressier, O., On the relation between truth maintenance and non-monotonic logics, Siemens Report INF 2 ARM-11-88.
18. Reiter, R., On closed world data bases, in Gallaire, H and Minker, J., editors, Logic and data base, pp.55-76, Plenum, New York, 1978.
19. Reiter, R., A logic for default reasoning, Artificial Intelligence 13, pp.81-132, 1980.
20. Shepherdson, J. C, Negation in logic programming, in Minker, J., editor, Foundations of deductive database and logic programming, pp.19-88, Morgan Kaufmann, Los Altos, 1987.
21. Van Gelder, A., Negation as failure using tight derivations for general logic programs, in Minker, J., editor, Foundations of deductive database and logic programming, pp.149-176, Morgan Kaufmann, Los Altos, 1987.