Constructing Probabilistic ATMS Using Extended Incidence Calculus

Citation for published version:
Liu, W & Bundy, A 1996, 'Constructing Probabilistic ATMS Using Extended Incidence Calculus', *International Journal of Approximate Reasoning: Uncertainty in Intelligent Systems*, vol. 15, no. 2. https://doi.org/10.1016/0888-613X(96)00031-X

Digital Object Identifier (DOI):
10.1016/0888-613X(96)00031-X

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
International Journal of Approximate Reasoning: Uncertainty in Intelligent Systems

Constructing Probabilistic ATMS Using Extended Incidence Calculus

Weiru Liu* **and Alan Bundy**
Department of Artificial Intelligence,
University of Edinburgh, Edinburgh EH1 1HN, UK.

May 12, 2003

Abstract

This paper discusses the relations between extended incidence calculus and assumption-based truth maintenance systems (ATMSs). We first prove that managing labels for statements (nodes) in an ATMS is equivalent to producing incidence sets of these statements in extended incidence calculus. We then demonstrate that the justification set for a node is functionally equivalent to the implication relation set for the same node in extended incidence calculus. As a consequence, extended incidence calculus can provide justifications for an ATMS, because implication relation sets are discovered by the system automatically. We also show that extended incidence calculus provides a theoretical basis for constructing a probabilistic ATMS by associating proper probability distributions on assumptions. In this way, we can not only produce labels for all nodes in the system, but also calculate the probability of any of such nodes in it. The nogood environments can also be obtained automatically. Therefore, extended incidence calculus and the ATMS are equivalent in carrying out inferences at both the symbolic level and the numerical level. This extends a result due to Laskey and Lehner.

KEYWORDS: incidence calculus, probabilistic reasoning, assumption based truth maintenance systems,

1 INTRODUCTION

One of the most important and difficult task of any intelligent system is to ultimately infer what we can confirm or possibly confirm given a set of facts. The set of statements we might confirm builds up our current beliefs towards the world we are concerning or we want to know. This set needs to be modified as more and more information is collected. Therefore our beliefs about the problem changes all the time and each of these changes is caused by some newly observed facts. To date, many approaches have been

[0] Address correspondence to Weiru Liu, School of Computing and Mathematics, University of Ulster at Jordanstown, Newtownabbey, Co.Antrim, BT37 0QB, UK

proposed to make such inference. Among them, the assumption-based truth maintenance system (ATMS) [6] provides an attractive mechanism to maintain and update the belief set.

The ATMS is a symbolic reasoning technique used in the artificial intelligence domain to deal with problems by providing dependent relations among statements during inference. This technique has been used in many areas such as fault diagnosis, trouble shooting, ... In practical applications, it has been found that a system using this technique can only infer results with absolutely true or false. It lacks the ability to draw plausible conclusions such as that a conclusion is true with a degree of belief. However in many cases, pieces of information from a knowledge base provide assumptions and premises with uncertainties. It is necessary to let the ATMS have the ability to cope with uncertainty problems.

In order to overcome this problem, some research on the integration of symbolic reasoning with numerical inference has been carried out to associate numerical uncertainties with ATMS [3], [4], [7], [10], [13], [14], [15], [16], [20], [21]. In [7], De Kleer and Williams use probability theory to deal with uncertainty associated with assumptions. In [10], [15], the authors use possibilistic logic to handle this problem. In [10], both assumptions and justifications are associated with uncertainty measures. The uncertainty values associated with justifications are used to select the path for deriving a node. Only those pathes with strong supporting relations are used to infer the corresponding nodes. [15] continues the work carried out in [10] and extends it to deal with a military data fusion application. [3], [4], [13], [16], [20], [21] all use Dempster-Shafer theory of evidence (DS theory) ([5], [22]) to calculate beliefs in statements. Among them, [16] studies a formal relation between DS theory and ATMS. It is proved in [16] that any belief network in DS theory can be translated into an ATMS structure. In such a system, inference is performed based on ATMS techniques while beliefs in statements are calculated by using probability theory.

One common limitation in all these extensions of the ATMS[1] is that the probabilities assigned to assumptions must be assumed probabilistically independent in order to calculate the degree of belief in a statement.

In this paper, we continue this research and intend to provide a general basis for constructing a probabilistic ATMS. The uncertainty technique we have chosen is extended incidence calculus. Incidence calculus was introduced in [1], [2] which aims at providing an automated reasoning technique to deal with uncertainty problems by associating classical propositional logic with probabilities. In [17], [18] this theory has been generalized considerablely to model a wider range of problems and the advanced theory is called *extended incidence calculus*. There are several reasons for us to choose extended incidence calculus to implement a probabilistic ATMS. First of all, apart from its numerical reasoning characters, extended incidence calculus also possesses some symbolic reasoning features. In extended incidence calculus, numerical uncertainties are not associateed with statements we want to infer, rather sets of possible worlds are associated with statements and uncertainties are associated with elements of possible worlds. Each possible world associated with a formula indicates that this formula is true under the support of this possible world. This is called the *indirect encoding of uncertainties*. In general, if we only consider the manipulation of incidence sets in incidence calculus, it is very similar to the calculation of labels of nodes in the ATMS. Secondly, as extended incidence calculus can calculate beliefs in statements after obtaining incidence sets, it can combine a numerical reasoning procedure and a symbolic reasoning procedure into

[1]Except the discussion in [10], [15] in which the topic was not discussed.

one mechanism. Finally, we have provided a more general combination technique in extended incidence calculus which can combine both dependent and independent pieces of information [17], [18]. So it is not necessary to assume the independence of probability distributions among assumptions as required in [3], [4], [7], [13], [16], [20], [21].

The main contributions of this paper are:

1. We prove that extended incidence calculus and the ATMS are equivalent at both the symbolic reasoning level (if we view the set of possible worlds in extended incidence calculus as the set of assumptions in an ATMS) and numerical inference level if we associate proper probabilistic distributions on assumptions. They can be translated into each other's form.

2. We show that the integration of symbolic and numerical reasoning patterns are possible and extended incidence calculus itself is a typical example of this unification. Extended incidence calculus can be regarded as a bridge between these two reasoning patterns.

3. in [17], [18] it has been proved that extended incidence calculus is equivalent to Dempster-Shafer theory of evidence in representing evidence and combining source-independent evidence. Therefore the result of investigating the relationship between extended incidence calculus and ATMS can provide a theoretical basis for some results in [16], namely the calculation of beliefs in nodes and the weight of conflict introduced by all evidence as well as its effect on individual nodes.

4. It is assumed that justifications must be supplied by the problem solver if one uses the ATMS techniques. We will show that extended incidence calculus can be used to provide justifications for nodes automatically without human's involvement. Therefore a complete automatic ATMS system is constructible.

5. The calculation of probabilities in nodes is done under the assumption that all given probability distributions are probabilistically independent. When this condition is not satisfied, the algorithm in [16] would not work. In [17], [18] we propose a more general combination mechanism to deal with the latter case. So extended incidence calculus can be used to help an ATMS to manage numerical uncertainties when it is necessary.

The paper is organized as follows. In the rest of this section, we will abstract the reasoning models in an ATMS and extended incidence calculus and then discuss their similarities. Section 2 introduces the basics of extended incidence calculus. In section 3 we introduce the ATMS notations and extend it by adding probabilities to assumptions. In section 4 we will show how to encode an ATMS structure into extended incidence calculus terminologies and perform the same inference in extended incidence calculus. We will explore how to manipulate labels of nodes and calculate degrees of belief in nodes in extended incidence calculus. In section 5 we will briefly discuss how to provide justifications from extended incidence calculus. In the concluding section, we summarize the paper.

1.1 The basic reasoning model in the ATMS

The truth maintenance system (TMS) [8] and later the ATMS [6] are both symbolic approaches to producing a set of statements in which we believe. The basic and central idea in such a system is that for each statement we believe in, a set of supporting statements (called labels or environments generally in the ATMS) is produced. A set of supporting statements is, in turn, obtained through a set of arguments attached to that statement (called justifications). In an ATMS, a justification of a statement (or called node) contains other statements (or nodes) from which the current statement can be derived. Justifications are specified by the system designer.

For instance, if we have two statements representing inference rules:

$$r_1 : p \rightarrow q$$

$$r_2 : q \rightarrow r$$

then logically we can infer that $r_3 : p \rightarrow r$. In an ATMS, if r_1, r_2 and r_3 are represented by $node_1$, $node_2$ and $node_3$ respectively, then $node_3$ is derivable from the conjunction of $node_1$ and $node_2$ and we call (r_1, r_2) a justification of $node_3$. Normally a rule may have several justifications. Furthermore if r_1 and r_2 are valid under the conditions that A and B are true respectively, then rule r_3 is valid under the condition that $A \wedge B$ is true, denoted as $\{A, B\}$. $\{A\}, \{B\}$ and $\{A, B\}$ are called sets of supporting statements (or environments) of r_1, r_2 and r_3 respectively. If we associate $node_3$ with the supporting statements such as $\{A, B\}$ and the dependent nodes such as $\{(r_1, r_2)\}$ then $node_3$ is generally in the form of

$$r_3 : p \rightarrow r, \{\{A, B\}...\}, \{(r_1, r_2)...\}$$

when $node_3$ has more than one justification. The collection of all possible sets of supporting environments is called the label of a node. If we use $L(r_3)$ to denote the label of $node_3$, then $\{A, B\} \in L(r_3)$. If we assume that r_1, r_2 hold without requiring any dependent relation on other nodes, then $node_1$ and $node_2$ are represented as

$$r_1 : p \rightarrow q, \{\{A\}\}, \{()\}$$

$$r_2 : q \rightarrow r, \{\{B\}\}, \{()\}$$

Therefore, we can infer a label for any node as long as its justifications are known.

The advantage of this reasoning mechanism is that the dependent and supporting relations among nodes are explicitly specified, in particular, the supporting relations among assumptions and other nodes. This is obviously useful when we want to retrieve the reasoning path. It is also helpful for belief revision.

The limitation of this reasoning pattern is that we cannot infer those statements which are probably true rather than absolutely true. However, if we attach numerical degrees of belief to the elements in the supporting set of a node, we may be able to infer a statement with a degree of belief. For example, if we know A is true with probability 0.8 and B is true with probability 0.7 and A and B are probabilistically independent, then the probability of $A \wedge B$ is true is 0.56. The belief in a node is considered as the probability of its label set. So for $node_3$, our belief in it is 0.56.

1.2 The basic reasoning model in extended incidence calculus

Incidence calculus was introduced by Bundy in [1], [2] to deal with problems in numerical reasoning. The special feature of this reasoning method is the indirect association of numerical uncertainty with formulae. In incidence calculus, probabilities are associated with the elements of a set of possible worlds (denoted as $\mathcal{W}$) and some formulae (called axioms) are associated with the subsets of the set of possible worlds. Each element in such a subset for a formula ϕ makes the formula true and this subset is normally called the incidence set of the formula, denoted as $i(\phi)$ ($i(\phi) \subseteq \mathcal{W}$). Our belief in a formula is regarded as the probability weight of its incidence set. Assume that the set of possible worlds is $\mathcal{W}$ and r_1, r_2 are two axioms in an incidence calculus theory and the incidence sets for r_1 and r_2 are $i(p \rightarrow q) = W_1$ and $i(q \rightarrow r) = W_2$, then the incidence set of $(p \rightarrow q \wedge q \rightarrow r)$ is $W_1 \cap W_2$. As formula $p \rightarrow r$ holds when formula $p \rightarrow q \wedge q \rightarrow r$ holds, the incidence set of $p \rightarrow q \wedge q \rightarrow r$ must be a subset of the incidence set of r_3. So $W_1 \cap W_2$ makes $p \rightarrow r$ true and $W_1 \cap W_2 \subseteq i(p \rightarrow r)$ true. So the propagation of incidences of formulaes is done through implication relations.

If $\mathcal{W}$ contains the set of assumptions in an ATMS, W_1, W_2 are the subsets of $\mathcal{W}$ and $W_1 \cap W_2$ is regarded as the conjunction of the elements in $W_1 \cap W_2$, then the manipulation of an incidence set is similar to the derivation of a label.

1.3 Similarities of the two reasoning models

Abstractly if we view the set of possible worlds in incidence calculus as the set of assumptions in an ATMS, and view the calculation of the incidence sets of formulae as the calculation of labels of nodes in the ATMS, then the two reasoning patterns are similar. Furthermore, as the probability weight of an incidence set can be calculated, incidence calculus has associated numerical uncertainty with symbolic reasoning into one mechanism. Incidence calculus has no such indications as justifications during its inference procedure. The implication relations are discovered automatically.

The apparent similarity of these two reasoning patterns motivates us to explore their relations more deeply. We focus our attention on the production of labels in the ATMS and calculations of incidence sets in incidence calculus. We will prove that the two reasoning mechanisms are equivalent in producing dependent relations among statements. As incidence calculus can draw a conclusion with a numerical degree of belief on it, incidence calculus actually possesses some features of both symbolic and numerical reasoning approaches. Therefore, incidence calculus can be used as a theoretical basis for the implementation of a probabilistic ATMS by providing both labels and degrees of belief of statements and as an automatic reasoning model to provide justifications for an ATMS.

2 EXTENDED INCIDENCE CALCULUS

2.1 Basics of extended incidence calculus

Incidence calculus [1], [2] starts with two sets, the set P contains propositions and the set $\mathcal{W}$ consists of possible worlds with a probability distribution on them. For each element w of $\mathcal{W}$, the probability on w, $\varrho(w)$, is known and $\Sigma\varrho(w) = 1$. From the set P, using logical operators $\wedge, \vee, \neg, \rightarrow$, a set of logical formulae are formed which is called the language set of P, denoted as $\mathcal{L}(P)$. The elements in the set $\mathcal{W}$ may make some formulae in $\mathcal{L}(P)$ true. For any $\phi \in \mathcal{L}(P)$, if every element in a subset W_1 of $\mathcal{W}$ makes ϕ true and W_1 is the maximal subset of this kind, then W_1 is represented as $i(\phi)$ in an incidence calculus theory and it is called the incidence set of ϕ. Therefore, the supporting set of a formula ϕ is $i(\phi)$ and its probability is $p(\phi) = wp(W_1)$ where $wp(W_1) = \Sigma_{w \in W_1} \varrho(w)$. It is assumed that $i(\perp) = \{\}$ and

$i(T) = \mathcal{W}$ where $\perp, T$ represent *false* and *true* respectively. In [17], [18] incidence calculus is extended in three aspects so that the advanced reasoning mechanism is more powerful. This advanced mechanism is called extended incidence calculus. In the following, we only introduce extended incidence calculus.

Definition 1: *Generalized incidence calculus theories*[2]

A generalized incidence calculus theory is a quintuple $< \mathcal{W}, \varrho, P, \mathcal{A}, i >$ *where* $\mathcal{W}$ *is a set of possible worlds with a probability distribution* ϱ, P *is a set of propositions and* $\mathcal{A}$ *is a subset of* $\mathcal{L}(P)$ *which is called a set of axioms. The function* i *assigns an incidence set to every formula in* $\mathcal{A}$. *For any two formulae in* $\mathcal{A}$, *we have*

$$i(\phi \wedge \psi) = i(\phi) \cap i(\psi)$$

Based on this definition, given two formulae $\phi, \psi \in \mathcal{A}$, we have $i(\phi) \subseteq i(\psi)$ if $\phi \models \psi$. For any other formula $\phi \in \mathcal{L}(P) \setminus \mathcal{A}$, it is possible to get the lower bound $i_*(\phi)$ of its incidence set as

$$i_*(\phi) = \bigcup_{\psi \models \phi} i(\psi) \tag{1}$$

For $\psi \in \mathcal{A}$, $\psi \models \phi$ means that formula $\psi \rightarrow \phi$ is valid (a tautology). The degree of our belief in a formula is defined as $p_*(\phi) = wp(i_*(\phi))$.

Definition 2: *Semantic implication set*

For any formula $\phi \in \mathcal{L}(P)$, *if* $\psi \models \phi$ *then* ϕ *is said to be semantically implied by* ψ, *denoted as* $\psi \models \phi$. *Let* $SI(\phi) = \{\psi \mid \psi \models \phi, \forall \psi \in \mathcal{A}\}$, *set* $SI(\phi)$ *is called a semantical implication set of* ϕ.

Definition 3: *Essential semantic implication set*

Furthermore, let $ESI(\phi)$ *be a subset of* $SI(\phi)$ *which satisfies the condition that a formula* ψ *is in* $ESI(\phi)$ *if for any* ψ' *in* $SI(\phi)$ *then* $\psi \not\models \psi'$, *then* $ESI(\phi)$ *is called an essential semantical implication set of* ϕ. *This is denoted as* $ESI(\phi) \models \phi$.

Proposition 1 *If* $ESI(\phi)$ *and* $ESI'(\phi)$ *are the two essential semantic implication sets of formula* ϕ *coming from the same generalized incidence calculus theory, then* $ESI(\phi) = ESI'(\phi)$.

Proof

Suppose that $ESI(\phi)$ and $ESI'(\phi)$ are different and further suppose that a formula ψ is in $ESI(\phi)$ but not in $ESI'(\phi)$. If $\psi \in ESI(\phi)$ then for any formula $\psi' \in SI(\phi)$, we have that $\psi \not\models \psi'$.

However, as $\psi \notin ESI'(\phi)$, then there is at least one formula ψ'' ($\psi'' \in SI(\phi)$) which makes the following equation true $\psi \models \psi''$. So according to Definition 3, $\psi \notin ESI(\phi)$. Conflict. Therefore, $ESI(\phi) = ESI'(\phi)$ and the essential semantic implication set is unique.

QED

[2]The original definition of incidence calculus theories in [2] is stricter than the definition here. More details on generalized incidence calculus theories can be found in [17], [18], [19].

It will be proved later that the essential semantic implication set of a formula is exactly the same as the set of justifications of that formula in an ATMS.

Example 1

Suppose we have a generalized incidence calculus theory and we know that the following five inference rules

$$r_1 : e \to d$$

$$r_2 : d \to b$$

$$r_3 : b \to a$$

$$r_4 : d \to c$$

$$r_5 : c \to a$$

are in the language set. Further suppose that the set of axioms $\mathcal{A}$ contains these five rules and all the possible conjunctions of them, then the lower bounds of incidence set of other formulae can be inferred. For instance, for formula $e \to a$, the lower bound of its incidence set is

$$i_*(e \to a) = \bigcup_{\phi \models (e \to a)} i(\phi)$$

According to Definition 2, all the formulae ϕ in $\mathcal{A}$ satisfying the condition that $\phi \models (e \to a)$ are in the semantic implication set. So the calculation of lower bounds of incidence sets can be restated as:

$$i_*(\psi) = \bigcup_{\phi \in SI(\psi)} i(\phi)$$

In this example, there are in total seven axioms satisfying this requirement, so there are seven axioms in $SI(e \to a)$.

$$
\begin{array}{l}
(e \to d) \wedge (d \to b) \wedge (b \to a) \\
(e \to d) \wedge (d \to c) \wedge (c \to a) \\
(e \to d) \wedge (d \to c) \wedge (c \to a) \wedge (d \to b) \\
(e \to d) \wedge (d \to c) \wedge (c \to a) \wedge (b \to a) \\
(e \to d) \wedge (d \to c) \wedge (c \to a) \wedge (d \to b) \wedge (b \to a) \\
(e \to d) \wedge (d \to b) \wedge (b \to a) \wedge (d \to c) \\
(e \to d) \wedge (d \to b) \wedge (b \to a) \wedge (c \to a)
\end{array}
$$

However if we examine these seven axioms closely, we will find that only the first two axioms are necessary to be considered if we want to get $i_*(e \to a)$. The rest are unnecessary as their incidence sets are included into the incidence sets of the first two axioms. Based on Definition 3, these two axioms are in the essential semantic implication set of $e \to a$ and this set only has these two axioms. Therefore the following proposition is natural.

Proposition 2 *If $SI(\phi)$ and $ESI(\phi)$ are a semantic implication set and an essential semantic implication set of ϕ, then the following equation holds:*

$$i_*(\phi) = i_*(SI(\phi)) = i_*(ESI(\phi))$$

where $i_*(SI(\phi)) = \bigcup_{\phi_j \in SI(\phi)} i(\phi_j)$.

PROOF

Assume a set of axioms in a generalized incidence calculus theory is $\mathcal{A}$. For a formula ϕ, when $\phi \in \mathcal{A}$, we have

$$\phi \in SI(\phi), \phi \in ESI(\phi), ESI = \{\phi\}$$

so

$$i_*(\phi) = i(\phi) = i_*(SI(\phi)) = i_*(ESI(\phi))$$

When $\phi \notin \mathcal{A}$, we have a set of formulae $\phi_1, ..., \phi_n \in \mathcal{A}$ ($n \geq 1$) each of which implies ϕ. So $SI(\phi) = \{\phi_1, ..., \phi_n\}$. Assume that the elements in $ESI(\phi)$ are $\psi_1, ..., \psi_m$, then for ψ_j, there will be some formulae $\phi_{j'}$ (at least ψ_j itself) in $SI(\phi)$ which make the following equation hold

$$\phi_{j'} \models \psi_j$$

Let SI_{ψ_j} be a set containing these $\phi_{j'}$, *i.e.* $SI_{\psi_j} = \{\phi_{j'} \mid \phi_{j'} \models \psi_j\}$, then we have $i_*(\psi_j) = i_*(SI_{\psi_j})$ because $i(\phi_{j'}) \subseteq i(\psi_j)$. Repeating this procedure for each formula in $ESI(\psi)$, we obtain the following equation

$$i_*(ESI(\phi)) = \cup_{\psi_j} i_*(SI_{\psi_j})$$

To prove

$$i_*(SI(\phi)) = i_*(ESI(\phi))$$

we need to prove that

$$i_*(SI(\phi)) = \cup_{\psi_j} i_*(SI_{\psi_j})$$

Assume that $i_*(SI(\phi)) \setminus \cup_{\psi_j} i_*(SI_{\psi_j}) = S \neq \{\}$, we have

$S \neq \{\}$ and $w \in S$
$\Rightarrow w \in i_*(SI(\phi)) \setminus \cup_{\psi_j} i_*(SI_{\psi_j})$
$\Rightarrow (\exists \varphi)\varphi \in SI(\phi), \varphi \notin ESI(\phi), w \in i(\varphi)$
$\Rightarrow (\exists \varphi')\varphi' \in SI(\phi), \varphi \models \varphi', \varphi' \notin ESI(\phi)$ (otherwise $\varphi \in SI_{\varphi'}$ and $\varphi \notin SI(\phi)$)
$\Rightarrow (\exists \varphi'')\varphi'' \in SI(\phi), \varphi' \models \varphi'', \varphi'' \notin ESI(\phi)$
$\Rightarrow$ ... (repeat this procedure until we find φ_t)
$\Rightarrow (\exists \varphi_t)\varphi_t \in SI(\phi), \varphi_{t-1} \models \varphi_t, \varphi_t \notin ESI(\phi)$ and $\nexists \varphi'_t, \varphi_t \models \varphi'_t$ (as $\mathcal{A}$ is finite)
$\Rightarrow \varphi_t \notin ESI(\phi)$ and $\varphi_t \in ESI(\phi)$

Conflict, so S is empty. Therefore, $i_*(SI(\phi)) = i_*(ESI(\phi))$ and $i_*(\phi) = i_*(SI(\phi))$.

END

Based on a generalized incidence calculus theory, the efficiency of calculating an incidence set for a formula is very much dependent on the speed of finding its semantic implication set as well as the essential semantic implication set.

2.2 Combining several generalized incidence calculus theories

An ATMS has the ability to make inferences based on more than one piece of information. In the following we will see how to deal with multiple pieces of information in extended incidence calculus in general.

Given a generalized incidence calculus theory, beliefs in formulae are derivable. Usually we consider that each generalized incidence calculus theory carries the information provided by one piece of evidence. If we have multiple pieces of evidence on a problem and their information is carried by multiple generalized incidence calculus theories, then we need to combine them in order to reach a conclusion from all the available information. The combination of multiple generalized incidence calculus theories is done using a combination rule in extended incidence calculus [16, 17]. Given two theories

$$< \mathcal{W}, \varrho, P, \mathcal{A}_1, i_1 >$$

$$< \mathcal{W}, \varrho, P, \mathcal{A}_2, i_2 >$$

the combination rule produces the third generalized incidence calculus theory as $< \mathcal{W} \setminus \mathcal{W}_0, \varrho', P, \mathcal{A}, i >$ where

$$\mathcal{W}_0 = \bigcup_{\phi \wedge \psi = \perp} i_1(\phi) \cap i_2(\psi) \qquad \phi \in \mathcal{A}_1, \psi \in \mathcal{A}_2$$

$$\mathcal{A} = \{\varphi \mid \varphi = \phi \wedge \psi, where\ \phi \in \mathcal{A}_1, \quad \psi \in \mathcal{A}_2, \quad \varphi \neq \perp\}$$

and

$$i(\varphi) = \bigcup_{(\phi \wedge \psi \models \varphi)} i_1(\phi) \cap i_2(\psi) \qquad \varphi \in \mathcal{A}$$

The probability distribution on $\mathcal{W} \setminus \mathcal{W}_0$ is updated as

$$\varrho'(w) = \frac{\varrho(w)}{1 - \Sigma_{w' \in \mathcal{W}_0} \varrho(w')} \qquad w \in \mathcal{W} \setminus \mathcal{W}_0$$

The special case of the rule is when two generalized incidence calculus theories are given on different sets of possible worlds and the two sets are probabilistically independent (or DS-Independent[3]), the combination can be performed using the Corollary 1 in [17]. Given that

$$< \mathcal{W}_1, \varrho_1, P, \mathcal{A}_1, i_1 >$$

$$< \mathcal{W}_2, \varrho_2, P, \mathcal{A}_2, i_2 >$$

[3]See definition and explanation in [18], [23]. In the analysis in [18], [23], two sets of possible worlds are probabilistically independent cannot guarantee they are DS-Independent when their common original source is known. In the case that original source is the set product of these two sets, their probabilistic independence also implies their DS-Independence. In this paper, as we only consider the latter case, we will use term *probabilistically independence* to name the DS-Independence among two sets.

applying Corollary 1 we get a combined theory $< \mathcal{W}_3, \varrho_3, P, \mathcal{A}_3, i_3 >$ where

$$\mathcal{W}_0 = \bigcup_{\phi \wedge \psi = \perp} i_1(\phi) \otimes i_2(\psi) \qquad \phi \in \mathcal{A}_1, \psi \in \mathcal{A}_2$$

$$\mathcal{W}_3 = \mathcal{W}_1 \otimes \mathcal{W}_2 \setminus \mathcal{W}_0$$

$$\varrho_3(w) = \varrho_3((w_{1i}, w_{2j})) = \frac{\varrho_1(w_{1i})\varrho_2(w_{2j})}{1 - \sum_{(w'_{1i}, w'_{2j}) \in \mathcal{W}_0} \varrho_1(w'_{1i})\varrho_2(w'_{2j})}$$

$$\mathcal{A}_3 = \mathcal{A} \text{ as defined above}$$

and

$$i_3(\varphi) = \bigcup_{\phi \wedge \psi \models \varphi} (i_1(\phi) \otimes i_2(\psi)) \setminus \mathcal{W}_0 \qquad \phi \in \mathcal{A}_1, \psi \in \mathcal{A}_2$$

In general a pair (w_{1i}, w_{2j}) is an element of $\mathcal{W}_1 \otimes \mathcal{W}_2 \setminus \mathcal{W}_0$. It is required that T is automatically added into a set of axioms $\mathcal{A}$ if $\cup_{\phi \in \mathcal{A}} i(\phi) \subset W$.

Similarly if there are several generalized incidence calculus theories and the corresponding probability spaces are probabilistically independent, the combined result will be $< \mathcal{W}, \varrho, P, \mathcal{A}, i >$. This result is also the same as that obtained by combining the theories one by one.

$$\mathcal{W}_0 = \bigcup_{\phi_1 \wedge \ldots \wedge \phi_n = \perp} i_1(\phi_1) \otimes \ldots \otimes i_n(\phi_n) \ \textit{where} \ \phi_i \in \mathcal{A}_i \tag{2}$$

$$\mathcal{W} = \mathcal{W}_1 \otimes \ldots \otimes \mathcal{W}_n \setminus \mathcal{W}_0$$

$$\varrho(w) = \varrho((w_{1i}, \ldots, w_{nj})) = \frac{\varrho_1(w_{1i}) \ldots \varrho_n(w_{nj})}{1 - \sum_{(w'_{1i}, \ldots, w'_{nj}) \in \mathcal{W}_0} \varrho_1(w'_{1i}) \ldots \varrho_n(w'_{nj})}$$

$$\mathcal{A} = \{\psi \mid \psi = \wedge \phi_j, \phi_j \in \mathcal{A}_j, \psi \neq \perp\}$$

and

$$i(\varphi) = \bigcup_{\phi_1 \wedge \ldots \wedge \phi_n \models \varphi} (i_1(\phi_1) \otimes \ldots \otimes i_n(\phi_n)) \setminus \mathcal{W}_0 \ \textit{where} \ \phi_i \in \mathcal{A}_i$$

Now we look at an example. Suppose that there are two generalized incidence calculus theories:

$$< \{X, \neg X\}, \varrho_1, P, \{d \rightarrow b, T\}, i_1(d \rightarrow b) = \{X\}, i_1(T) = \{X, \neg X\} >$$

$$< \{V, \neg V\}, \varrho_2, P, \{b \rightarrow a, T\}, i_2(b \rightarrow a) = \{V\}, i_2(T) = \{V, \neg V\} >$$

if the two sets of possible worlds are probabilistically independent, then using the above corollary the combined theory is

$$< S_X \otimes S_V, \varrho_3, P, \{d \rightarrow b \wedge b \rightarrow a, \ldots, T\}, i_3(d \rightarrow b \wedge b \rightarrow a) = \{(X, V)\} \ldots >$$

Table 1 below shows the combination procedure.

ϕ $i(\phi)$	$d \to b$ $\{X\}$	T $\{X, \neg X\}$
$b \to a$ $\{V\}$	$d \to b \wedge b \to a$ $\{(X, V)\}$	$b \to a$ $\{X, \neg X\} \otimes \{V\}$
T $\{V, \neg V\}$	$d \to b$ $\{X\} \otimes \{V, \neg V\}$	T $\{X, \neg X\} \otimes \{V, \neg V\}$

Table 1. Combination of two independent generalized incidence calculus theories

The first two rows in the table represent the first generalized incidence calculus theory and the first column represent the second theory. From the combined theory, we have $i_*(d \to a) = \bigcup_{\phi \models (d \to a)} i(\phi) = \{(X, V)\}$. If we know that $\varrho_1(X) = 0.75, \varrho_2(V) = 0.8$, then $p_*(d \to a) = 0.6$ which is our belief in formula $d \to a$. In this case, the conflict set $\mathcal{W}_0$ is empty.

3 EXTENDING ASSUMPTION-BASED TRUTH MAINTENANCE SYSTEMS

The ATMS was introduced by de Kleer [6] based on the TMS [8] in which a special set of arguments, named as assumptions, are particularly addressed. Considering an inference rule $a \to b$, normally in propositional logic this rule tells us that if a is observed then b is believed to be true. In this procedure the information supporting the inference from a to b is assumed to be true by default. If this information is supplemented then the rule can be written as:

$$a \wedge C \to b$$

where C is regarded as the information related to the rule but hidden behind the rule. In an expert system, C can be thought of the rule strength m. While in an ATMS, C is called an assumption[4]. In the absence of information, assumptions are assumed to be true in the procedure of carrying out inferences. When a conflict is discovered, some of the assumptions will be assigned *false* to prevent the firing of relevant rules. In this section, we extend the ATMS by associating probabilities on assumptions in order to establish formal and theoretical relations between a probabilistic ATMS and incidence calculus.

3.1 Non-redundant justification sets and environments

We briefly describe the ATMS below.

node: a node (called a problem-solver's datum) in an ATMS represents any datum unit used in the system. This datum unit can be a proposition or any formula in the propositional language which the system uses. The truth and falsity of a datum unit is inferred during the system processing procedure.

[4]We follow de Kleer's convention that upper-case letters are used to represent assumptions.

assumptions: a set of distinguished nodes which are believed to be true without requiring any preconditions are called assumptions.

justifications: justifications are supplied by the problem-solver. A justification for a node contains those nodes from which it can be derived. Usually, a node has several justifications representing multiple paths to infer the node.

label: a set of assumptions is called *an environment* of a node if the node holds under this environment. The label of a node contains all collections of such environments. Each environment in a label consists of non-redundant assumptions.

nogood: there is a nogood node in an ATMS system, any environment in which falsity is derived is included in the label of nogood.

In an ATMS, each node is associated with a label and a set of justifications and the node is normally denoted as

$$< node_j, label, justifications >$$

The inference procedure in the ATMS propagates assumptions along justifications.

Both the label and the justifications for a node can be explained as material implications. Given a node c with label $\{\{A_1, A_2, ...\} \quad \{B_1, B_2, ...\}...\}$ and with justifications $\{(z_1,\ z_2,\ ...)\ (y_1,\ y_2,\ ...)...\}$, the meaning of the label of c is that the conjunction of assumptions in each environment makes c true, such as $A_1 \wedge A_2 ...$ of environment $\{A_1, A_2 ...\}$ makes c true. So $L(c)$ is a set containing conjunctions of assumptions. $L(c) = \{(A_1 \wedge A_2 \wedge ...), (B_1 \wedge B_2 \wedge ...)...\}$. The following relation is true.

$$(A_1 \wedge A_2 \wedge ...) \vee (B_1 \wedge B_2 \wedge ...) \vee ... \rightarrow c$$

Each justification of c also represents an implication, that is, for justification $(z_1, z_2, ...)$, if $z_1 \wedge z_2 \wedge ...$ is proved to be true, then c is true as well. So there is similar implication relation:

$$(z_1 \wedge z_2 \wedge ...) \vee (y_1 \wedge y_2 \wedge ...) \vee ... \rightarrow c$$

The relations between a justification and its node states that the conjunction of $z_i(y_j)$ logically supports the conclusion c. If we consider z_i and c as formulae in a propositional language, then $\wedge_i z_i$ is a formula in the language which implies c, that is, formula $\wedge_i z_i \rightarrow c$ is always true. In general if we let $j(c) = \{(z_1 \wedge z_2 \wedge ...), (y_1 \wedge y_2 \wedge ...)...\}$, then every element in $j(c)$ semantically implies c, so $j(c) \models c$.

In general each justification is nonredundant. That is, deleting any element in an justification will destroy the implication relation of this justification to its node. For any two justifications for one node, usually these two justifications don't imply each other. That is one justification cannot be inferred from another. If one justification can be inferred by another, then the effect of this justification will be covered by the latter one. The same rules also apply to the environments for a node. So any environment is nonredundant and any two environments of a node have at lease one different assumption. We will show this in the following example.

Example 2

The five inference rules given at the beginning of Example 1 can be encoded into a set of ATMS nodes as[5]

$$\begin{aligned}
&node_1 :< e \to d, \{\{Z\}\}, \{(Z)\} > \\
&node_2 :< d \to b, \{\{X\}\}, \{(X)\} > \\
&node_3 :< b \to a, \{\{V\}\}, \{(V)\} > \\
&node_4 :< d \to c, \{\{Y\}\}, \{(Y)\} > \\
&node_5 :< c \to a, \{\{W\}\}, \{(W)\} >
\end{aligned}$$

Similarly we encode another two inference rules in this ATMS as

$$\begin{aligned}
&node_6 :< d \to a, \{\{X, V\}, \{Y, W\}\}, \{(node_2, node_3), (node_4, node_5)\} > \\
&node_7 :< e \to a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \{(node_1, node_6)\}
\end{aligned}$$

or replacing $node_6$ by its justification set

$$\begin{aligned}
&node_7 :< e \to a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \\
&\quad \{(node_1, node_2, node_3), (node_1, node_4, node_5)\}
\end{aligned}$$

We should notice that $(node_1, node_2, node_3)$ also implies $node_6$, but it is not in the justification set of $node_6$ as the effect of this justification has been covered by the justification $(node_2, node_3)$. The same thing happens to $node_7$ as well.

In fact there are in total seven conjunctions of nodes make $node_7$ true, but only two of them are included in the justification set of $node_7$. These seven conjunctions of nodes and the two of them used in the justification set are exactly the same as the *semantic implication set* and the *essential semantic implication set* for formula $e \to a$ in extended incidence calculus (see Example 1). If the essential semantic implication set of a formula is known, then this set can be used as justifications for the node. That is why we use extended incidence calculus to provide justifications for nodes. We will discuss this in detail in section 5.3.

The justification set of a node in an ATMS contains implication relations among a set of nodes and this desired node. If we require that a justification set of a node is non-redundant, then deleting any justification from the justification set of a node will cut off a path which can derive the node. From any given justification set, we can always get a non-redundant justification set from it and these two sets give out the same environments. For any inference chain which can derive the node, there must exist a justification. This justification contains fewer nodes then the chain but can infer the same result. The labels of nodes are also non-redundant. The non-redundancy of a label means either that for any two environments in the label of a node, one environment cannot be inferred from another or that deleting any assumption (or assumptions) in an environment will destroy the supporting relation among this node and the environment.

For $node_7$, the non-redundant justification set and label are

$$\{(node_1, node_2, node_3)$$

$$(node_1, node_4, node_5)\}$$

[5] A node with only an assumption (or assumptions) in both its label and its justifications means that this node is supported and dependent on this assumption (or assumptions) only.

and

$$\{\{Z, X, V\}, \{Z, Y, W\}\}$$

respectively.

3.2 Probabilistic assumption sets

In an ATMS, all nodes can be divided into four types: *assumptions, assumed nodes, premises,* and *derived nodes.* An assumption node is a node whose label contains a singleton environment mentioning itself, such as $< A, \{\{A\}\}, \{(A)\} >$.

An *assumed node* is a node which has justifications mentioning only assumptions[6]. For instance $< a, \{\{A\}\}, \{(A)\} >$ or $< b, \{\{A, B\}\}, \{(A, B)\} >$. All other nodes are either premises or derived. A *premise* (or a fact) has an empty justification and empty label set, *i.e.*, it holds without any preconditions. A *derived node* usually doesn't include assumptions in its justifications, such as $\{c, \{\{A, B\}\}, \{(a, b)\}\}$. In general, if we keep the restriction that non-assumptions cannot become assumptions, or assumptions cannot become another type of node [6], then it is possible to keep all assumptions in one set and other nodes in another set, and the two sets are distinct.

The inference result of a node has one of three values: Believed, Disbelieved and Unknown. If one of the environments in the label c is believed, then c is believed. If one of the environments in the label $\neg c$ is believed, then c is disbelieved, otherwise c is unknown. When both c and $\neg c$ are believed, there is a conflict and falsity is derived. In this case, some of the previous results should be retrieved and reinferred, *e.g.*, delete *nogood* environments from those labels of nodes where they appear. Such kinds of inference in an ATMS produce only three possible values. It cannot represent a plausible conclusion d with a degree of belief. Attempts to attach uncertain numbers with assumptions in the ATMS have appeared in [3], [4], [7], [10], [15], [16]. The belief of a node is identified as the probability of its label $Bel(c) = Pr(L(c))$.

For example [20], the rule *Turn the key* $\rightarrow$ *start the engine with* 0.8 can be represented in the ATMS as

$$< b \rightarrow a, \{\{B\}\}, \{(B)\} >$$

where B stands for an assumption (or a set of assumptions) which supports the implication relation $b \rightarrow a$ and assign 0.8 as the probability of B. a and b represent propositions *'start the engine'* and *'turn the key'* respectively.

Assume that for node b we have $< b, \{\{A\}\}, \{(A)\} >$, then the justification for node a is $b \wedge (b \rightarrow a) \Rightarrow a$. That is for node a we have $< a, \{\{A, B\}\}, \{(b, b \rightarrow a)\} >$. a is a derived node.

Therefore $Bel(a) = Pr(L(a)) = Pr(A \wedge B) = 0.8$, if the probability distributions are probabilistically independent and the action *'turn the key'* is true, i.e., $p(A) = 1$.

In this way, principlely the ATMS has the ability to make plausible inferences with beliefs. For a simple case like this, the calculation of probabilities on nodes is not difficult to carry out. However, in most cases labels of nodes are very complicated and probability distributions on assumptions maybe somehow related. In these circumstances, calculating probabilities of labels of nodes is quite troublesome as shown in [16], [20]. We introduce the following two definitions to cope with this difficulty in general.

[6] In [7], an assumed node has only one justification mentioning one assumption.

Definition 4: Probabilistic assumption set[7]

A set $\{A_1, ..., A_n\}$, *denoted as* $S_{A_1,...,A_n}$, *is called a probabilistic assumption set for assumptions* $A_1, ..., A_n$ *if the probabilities on* $A_1, ..., A_n$ *are given by a probability distribution p from a piece of evidence and* $\Sigma_{D\in\{A_1,...,A_n\}} p(D) = 1$. *The simplest probabilistic assumption set has two elements* A *and* $\neg A$, *denoted as* $S_{A,\neg A}$. *For any two elements in a probabilistic assumption set, it is assumed that* $A_i \wedge A_j \Rightarrow \perp$. *For all elements in the set, we have* $\vee_j A_j = true$ *for* $j = 1, ..., n$.

For two distinct probabilistic assumption sets $S_{A_1,...,A_n}$ and $S_{B_1,...B_m}$, the unified probabilistic assumption set is defined as $S_{A_1,...,A_n,B_1,...B_m} = S_{A_1,...,A_n} \otimes S_{B_1,...,B_m} = \{(A_i, B_j) \mid A_i \in S_{A_1,...,A_n}, B_j \in S_{B_1,...,B_m}\}$ where $\otimes$ means set product and $p(A_i, B_j) = p_1(A_i) \times p_2(B_j)$. p_1 and p_2 are the probability distributions on $S_{A_1,...,A_n}$ and $S_{B_1,...,B_m}$, respectively.

Example 3

Assume that the five assumptions in Example 2 are in different probabilistic assumption sets. An environment for $node_6$ derived from justification $\{(node_2, node_3)\}$ is $\{\{X, V\}\}$, then the joint probabilistic assumption set for this environment is $S_{X,\neg X} \otimes S_{V,\neg V}$. Similarly the joint probabilistic assumption set for environment $\{\{Y, W\}\}$ is $S_{Y,\neg Y} \otimes S_{W,\neg W}$.

Definition 5: *Full extension of a label*

Assume that an environment of a node n is $\{A, B, ..., C\}$ *where* $A, B, ..., C$ *are in different probabilistic assumption sets* $S_{A_1,...,A_x}, S_{B_1,...,B_y}$ *and* $S_{C_1,...,C_z}$. *Because* $A \wedge B \wedge ... \wedge C = A \wedge B \wedge ... \wedge C \wedge (\vee E_j \mid E_j \in S_{E_1,...,E_t})$, $A \wedge B \wedge ... \wedge C \rightarrow n$ *and* $A \wedge B \wedge ... \wedge C \wedge (\vee_j E_j \mid E_j \in S_{E_1,...E_t}) \rightarrow n$ *are all true (where* $S_{E_1,...,E_t}$ *is a probabilistic assumption set which is different from* $S_{A_1,...A_x}, S_{B_1,...,B_y}$ *and* $S_{C_1,...,C_z}$*).* $\{A, B, ..., C\} \otimes S_{E_1,...,E_t}$ *is called a full extension of the environment to* $S_{E_1,...E_t}$. *If there are in total m probabilistic assumption sets in the ATMS, then* $\{A, B, ..., C\} \otimes S_{E_1,...,E_t} \otimes ... \otimes S_{F_1,...,F_j}$ *is called the full extension of the environment to all assumptions, or simply called the full extension of the environment. Similarly if every environment in a label has been fully extended to all assumptions, then we call the result the full extension of the label, denoted as* $FL(n)$.

To understand the idea behind this definition, we look at Example 2 again. There are 5 probabilistic assumption sets in this ATMS structure, $S_{Z,\neg Z}$, $S_{X,\neg X}$, $S_{V,\neg V}$, $S_{Y,\neg Y}$ and $S_{W,\neg W}$. One environment of $node_6$ is $\{X, V\}$ which contains assumptions in two probabilistic assumption sets $S_{X,\neg X}$ and $S_{Y,\neg Y}$. Based on Definition 5 the full extension of this environment is

$$\{X, V\} \otimes S_{Z,\neg Z} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W}$$

and the full extension of label $L(node)$ is

$$\{X, V\} \otimes S_{Z,\neg Z} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W} \cup \{Y, W\} \otimes S_{X,\neg X} \otimes S_{V,\neg V} \otimes S_{Z,\neg Z}$$

Similarly, we are able to calculate full extensions for all environments of nodes.

In particular, let $L(\perp)$ represent all inconsistent environments (*i.e. nogood*) and let $FL(\perp)$ represent the full extension of them. If a label of a node is $L(c) = \{\{A_1, A_2, ...\}, \{B_1, B_2, ...\}, ...\}$, it means that

[7]Similar definition is given in [16] called an auxiliary hypothesis set.

$(A_1 \wedge A_2 \wedge ...) \vee (B_1 \wedge B_2 \wedge ...) \vee ... \to c$ is true. After we get the full extension of the label and represent it in disjunctive normal form (a disjunction of conjunctions), we have that $(A_1 \wedge A_2 \wedge ... \wedge B_1 \wedge ...C_1) \vee ... \vee (A_1 \wedge A_2 \wedge ...B_n \wedge ...C_1 \wedge ...) \vee ...(A_1 \wedge A_2 \wedge ... \wedge B_n \wedge ... \wedge C_m) \to c$ is true, each conjunction in the full extension contains the elements from different probabilistic assumption sets and any two such conjunctions are different. Such a full extension is convenient for calculating uncertainties related to assumptions.

The motivation of this definition comes from two aspects. First of all, although Laskey and Lehner have the definition of probabilistic assumption sets in [16] implicitly and give an algorithm to calculate the probability of a node based on its label, we are not satisfied with the algorithm they give. It lacks theoretical notation. Secondly, if we organize different assumptions into different probabilistic sets, we'd better adopt some set operations to deal with them. In this sense, the management method on sets of possible worlds in extended incidence calculus seems reasonable to be used here. These two reasons suggested us to give the above definition about how to extend a label into its full length notation and such a full extension is convenient for calculating uncertainties related to assumptions.

Example 4

In Example 3, we have two different probabilistic assumption sets for two environments of $node_6$. However the probability of $node_6$ cannot be obtained by calculating them separately and then adding them together. Doing so may over count the joint part in these two sets. The solution to this is to apply Definition 5 to each of these environments and we have full extensions for these two environments as

$$S_{Z,\neg Z} \otimes \{X, V\} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W}$$

$$S_{Z,\neg Z} \otimes S_{X,\neg X} \otimes S_{V,\neg V} \otimes \{Y, W\}$$

The full extension of the label of $node_6$ is the union of these two sets.

$$(S_{Z,\neg Z} \otimes \{X, V\} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W}) \cup (S_{Z,\neg Z} \otimes S_{X,\neg X} \otimes S_{V,\neg V} \otimes \{Y, W\})$$

or

$$S_{Z,\neg Z} \otimes (\{X, V\} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W} \cup S_{X,\neg X} \otimes S_{V,\neg V} \otimes \{Y, W\})$$

If we use p_Z to represent the probability distribution on probabilistic assumption set $S_{Z,\neg Z}$, then belief in this node is

$$\begin{aligned}
&Bel(node_6)\\
&= p_Z(S_{Z,\neg Z})(p_X(X)p_V(V)p_Y(S_{Y,\neg Y})p_W(S_{W,\neg W})+\\
&p_X(S_{X,\neg X})p_V(S_{V,\neg V})p_Y(Y)p_W(W)\ -p_X(X)p_V(V)p_Y(Y)p_W(W))\\
&= p_Z(S_{Z,\neg Z})(p_X(X)p_V(V) + p_Y(Y)p_W(W) - p_X(X)p_V(V)p_Y(Y)p_W(W))
\end{aligned}$$

In general if the *nogood* environments are not empty, those non-empty environments should be deleted from the label of a node. The probability of a node is then changed to:

$$Bel(node) = Pr(FL(a) \setminus FL(\perp))$$

4 CONSTRUCTING LABELS AND CALCULATING BELIEFS IN NODES USING EXTENDED INCIDENCE CALCULUS

We have introduced extended incidence calculus and the ATMS in the previous two sections. In this section we are going to draw some mapping relations among the components in these two reasoning mechanisms. Imagine that the joint set of set products of different probabilistic assumption sets in an ATMS corresponds to the set of possible worlds in a generalized incidence calculus theory and also imagine that the set of nodes (except assumptions) in an ATMS is translated into the language set $\mathcal{L}(P)$ of a suitable proposition set P in extended incidence calculus, then the supporting relation between the labels (which contain assumptions) and the set of nodes in the ATMS is similar to the supporting relation between the set of possible worlds and the language set in extended incidence calculus. This is the intuition behind our formal manipulation procedure for producing incidence sets (or the lower bounds) for formulae which can then be used to obtain labels for nodes in the ATMS.

4.1 An example

Now we will use an example (from [16]) to show how to manage assumptions in the ATMS in the way we manage sets of possible worlds in extended incidence calculus. We will solve this problem using ATMS techniques and extended incidence calculus respectively. The result shows that both inference mechanisms can be used to solve the same problem and the results are the same. It also indicates the procedure of transforming an ATMS into extended incidence calculus.

Example 5

Assume that we have five inference rules from Example 2 and fact e is observed, we want to infer our belief in other statements, such as a. This is shown in figure 1.

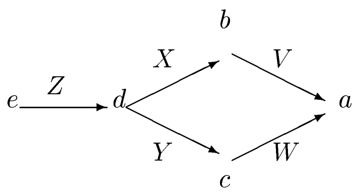

Figure 1. Semantic Network of Inference Rules

Approach 1: Solving this problem in an ATMS.

Assume that there are following nodes in an ATMS:
assumed nodes:

$$n_1 :< e \rightarrow d, \{\{Z\}\}, \{(Z)\} >$$

$$n_2 :< d \rightarrow b, \{\{X\}\}, \{(X)\} >$$

$$n_3 :< b \rightarrow a, \{\{V\}\}, \{(V)\} >$$

$$n_4 :< d \rightarrow c, \{\{Y\}\}, \{(Y)\} >$$

$$n_5 :< c \rightarrow a, \{\{W\}\}, \{(W)\} >$$

premise node:

$$n_8 :< e, \{\{\}\}, \{()\} >$$

derived nodes:

$$n_6 :< d \rightarrow a, \{\{X, V\}, \{Y, W\}\}, \{(n_2, n_3), (n_4, n_5)\} >$$
$$n_7 :< e \rightarrow a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \{(n_1, n_6)\} >$$

or replacing n_6 by its own justifications

$$n_7 :< e \rightarrow a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \{(n_1, n_2, n_3), (n_1, n_4, n_5\} >$$
$$n_9 :< a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \{(n_7, n_8)\} >$$

or

$$n_9 :< a, \{\{Z, X, V\}, \{Z, Y, W\}\}, \{(n_1, n_2, n_3, n_8), (n_1, n_4, n_5, n_8)\} >$$

assumption nodes: $< X, \{\{X\}\}, \{(X)\} >$ and so on.

It is not enough to know labels only if we are interested in calculating beliefs on nodes [20], [16]. We would have to manipulate labels in some way in order to get the beliefs. In our approach, we need to obtain the full extension of a label first. In order to do so, probabilistic assumption sets are required and some new assumptions need to be created when necessary. For the premise node e, if we associate it with a distinct assumption E, then node n'_8 can be rewritten as $n'_8 :< e, \{\{E\}\}, \{(E)\} >$. There are in total six probabilistic assumption sets. They are $S_{V,\neg V}$, $S_{W,\neg W}$, $S_{X,\neg X}$, $S_{Y,\neg Y}$, $S_{Z,\neg Z}$, $S_{E,\neg E}$.

The labels of derived nodes are obtained based on the justifications given by the problem solver, premise nodes and assumed nodes. The label of proposition a is $L(a) = \{\{Z, X, V\}\{Z, Y, W\}\}$ and its full extension is

$$FL(a) = S_{E,\neg E} \otimes \{Z\} \otimes (\{X, V\} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W} \cup S_{X,\neg X} \otimes S_{V,\neg V} \otimes \{Y, W\})$$

If we assume that different probability distributions on different assumption set are

$$p_V(V) = .7$$

$$p_W(W) = .8$$

$$p_X(X) = .6$$

$$p_Y(Y) = .75$$

$$p_Z(Z) = .8$$

$$p_E(E) = 1$$

and they are probabilistically independent, then the belief in node a is

$$Bel(a) = Pr(FL(a)) = 1 \times .8 \times (.6 \times .7 + .75 \times .8 - .6 \times .7 \times .75 \times .8) = 0.6144$$

A different calculation procedure can also be found in [16] which produces the same result.

Approach 2: Using extended incidence calculus to solve the problem.

Now let us see how his problem can be solved in extended incidence calculus. Suppose that we have the following six generalized incidence calculus theories

$$< S_{V,\neg V}, \varrho_1, P, \{b \to a, T\}, i_1(b \to a) = \{V\}, i_1(T) = S_{V,\neg V} >$$

$$< S_{W,\neg W}, \varrho_2, P, \{c \to a, T\}, i_2(c \to a) = \{W\}, i_2(T) = S_{W,\neg W} >$$

$$< S_{X,\neg X}, \varrho_3, P, \{d \to b, T\}, i_3(d \to b) = \{X\}, i_3(T) = S_{X,\neg X} >$$

$$< S_{Y,\neg Y}, \varrho_4, P, \{d \to c, T\}, i_4(d \to c) = \{Y\}, i_4(T) = S_{Y,\neg Y} >$$

$$< S_{Z,\neg Z}, \varrho_5, P, \{e \to d, T\}, i_5(e \to d) = \{Z\}, i_5(T) = S_{Z,\neg Z} >$$

$$< S_{E,\neg E}, \varrho_6(E) = 1, P, \{e\}, i_6(e) = \{E\}, i_6(T) = S_{E,\neg E} >$$

where $S_{V,\neg V}$, ..., $S_{Z,\neg Z}$, and $S_{E,\neg E}$ are probabilistic assumption sets.

As we assumed that sets of $S_{X,\neg X}, ..., S_{E,\neg E}$ are probabilistically independent, the combination of the first five theories produces a generalized incidence calculus theory $< S_7, \varrho_7, P, \mathcal{A}_7, i_7 >$ in which the joint set is $S_7 = S_{Z,\neg Z} \otimes S_{X,\neg X} \otimes S_{V,\neg V} \otimes S_{Y,\neg Y} \otimes S_{W,\neg W}$.

$i_7(d \to b \wedge b \to a) = S_{Z,\neg Z}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W}$
$\quad = S_{Z,\neg Z}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W}$
$i_7(d \to c \wedge c \to a) = S_{Z,\neg Z}\{Y\}\{W\}S_{X,\neg X}S_{V,\neg V}$
$i_7(d \to b \wedge b \to a \wedge d \to c \wedge c \to a) = S_{Z,\neg Z}\{X\}\{V\}\{Y\}\{W\}$
$i_7(e \to d \wedge d \to b \wedge b \to a) = \{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W}$
$i_7(e \to d \wedge d \to c \wedge c \to a) = \{Z\}\{Y\}\{W\}S_{X,\neg X}S_{V,\neg V}$

If we let $e \to d \wedge d \to b \wedge b \to a = \phi_1$ and $e \to d \wedge d \to c \wedge c \to a = \phi_2$, then

$$i_7(\phi_1 \wedge \phi_2) = \{Z\}\{X\}\{V\}\{Y\}\{W\}$$

Combining this theory with the sixth generalized incidence calculus theory we obtain
$i(e \wedge \phi_1) = S_{E,\neg E}\{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W}$
$i(e \wedge \phi_2) = S_{E,\neg E}\{Z\}\{Y\}\{W\}S_{X,\neg X}S_{V,\neg V}$
$i(e \wedge \phi_1 \wedge \phi_2) = S_{E,\neg E}\{Z\}\{X\}\{V\}\{Y\}\{W\}$. Because $e \wedge \phi_1 \to a$, $e \wedge \phi_2 \to a$ and $e \wedge \phi_1 \wedge \phi_2 \to a$, the following equation holds:

$$i_*(a) = i(e \wedge \phi_1) \cup i(e \wedge \phi_2) \cup i(e \wedge \phi_1 \wedge \phi_2)$$

$$= S_{E,\neg E}\{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{E,\neg E}S_{X,\neg X}S_{V,\neg V}\{Z\}\{Y\}\{W\}$$

and

$p_*(a) = wp(i_*(a))$
$\quad = wp(S_{E,\neg E}\{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{E,\neg E}S_{X,\neg X}S_{V,\neg V}\{Z\}\{Y\}\{W\})$
$\quad = wp(S_E) \times wp(\{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{X,\neg X}S_{V,\neg V}\{Z\}\{Y\}\{W\})$
$\quad = wp(S_{E,\neg E}) \times wp(\{Z\}) \times wp(\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{X,\neg X}S_{V,\neg V}\{Y\}\{W\})$
$\quad = wp(S_{E,\neg E}) \times wp(\{Z\}) \times (wp(\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W})$

$+wp(S_{X,\neg X}S_{V,\neg V}\{Y\}\{W\})\ -wp(\{X\}\{V\}\{Y\}\{W\}))$
$= 1 \times 0.8 \times (.6 \times .7 \times 1 \times 1 + 1 \times 1 \times .75 \times .8 - .6 \times .7 \times .75 \times .8) = 0.6144$

So our belief in a is also 0.6144.

Similarly we can obtain $i_*(d \to a)$, $i_*(e \to a)$ as:

$$i_*(d \to a) = S_{E,\neg E}S_{Z,\neg Z}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{E,\neg E}S_{Z,\neg Z}\{Y\}\{W\}S_{X,\neg X}S_{V,\neg V}$$

$$i_*(e \to a) = S_{E,\neg E}\{Z\}\{X\}\{V\}S_{Y,\neg Y}S_{W,\neg W} \cup S_{E,\neg E}\{Z\}\{Y\}\{W\}S_{X,\neg X}S_{V,\neg V}$$

These six generalized incidence calculus theories are in fact produced from assumed and premise nodes in the ATMS.

If we compare the full extensions of nodes in the ATMS and the lower bounds of incidence sets on formulae, we can find that the following equations hold:

$$i_*(d \to a) \equiv FL(d \to a) \qquad i_*(e \to a) \equiv FL(e \to a) \qquad i_*(a) \equiv FL(a)$$

That is the full extension of a node is the same as the lower bound of incidence set of the corresponding formula.

Here the symbol $\equiv$ is read as "equivalent to". An incidence set of a formula (or its lower bound) is equivalent to the full extension of the label of a node means that for an element $(a_1, a_2, ..., a_k)$ in the incidence set, the element $(a_1 \wedge a_2 \wedge ... \wedge a_k)$ is in $FL(*)$. In the following we give the general procedure of encoding a list of ATMS nodes by the equivalent generalized incidence calculus theories.

4.2 The algorithm of equivalent transformation from an ATMS into extended incidence calculus

Definition 6: Equivalent transformation algorithm

Given an ATMS we follow the following steps to convert it into generalized incidence calculus theories.

Step 1: divide the list of nodes into four sets: a set of assumption nodes, a set of assumed nodes, a set of derived nodes and a set of premises. The set of assumption nodes is called lower level nodes and the last three sets together are called higher level nodes. Based on the higher level nodes, a set of propositions P is established. A higher level node is either a proposition in P or a formula in $\mathcal{L}(P)$.

Step 2: from the set of assumption nodes, we can form a list of probabilistic assumption sets $S_{A_1,...,A_m}, S_{B_1,...B_n}, ...$, based on Definition 4. It is also assumed that these sets are probabilistically independent. If they are not independent, an extended ATMS cannot solve them.

Step 3: divide those assumed nodes into groups. If both node n_i and n_j are in group i, then n_i and n_j must satisfy one of the conditions: there exists an assumption A which is in an environment of $L(n_i)$ and also in an environment of $L(n_j)$ or an assumption in $L(n_i)$ and an assumption in $L(n_j)$ are in the same probabilistic assumption set. If n_i and n_j are in the same group and n_j and n_l are in the same group, then n_i, n_j and n_l should be in the same group.

Step 4: for any group k, create a corresponding structure $< \mathcal{W}_k, \varrho_k, P, i_k, \mathcal{A} >$. The set of axioms $\mathcal{A}$ contains assumed nodes in this group and all the possible conjunctions of them. The set of possible worlds $\mathcal{W}_k$ is either a probabilistic assumption set or the set product of several such sets if there is more than one probabilistic assumption set involved in the labels of these assumed nodes. For instance, if the label of node n_i is $\{\{A\},\{B\}\}$ and $S_{A_1,\ldots,A_m}, S_{B_1,\ldots,B_n}$ are different, then the set of possible worlds $\mathcal{W}_k$ should be $\mathcal{W}_k = S_{A_1,\ldots,A_m} \otimes S_{B_1,\ldots,B_n}$. The incidence function i_k is defined as $i_k(n_t) = L(n_i)$ and $i_k(n_t \wedge n_j) = L(n_t) \cap L(n_j)$. So i_k defined on $\mathcal{A}$ is closed under $\wedge$. We further define $i_k(false) = \{\}$ and $i_k(true) = \mathcal{W}_k$, then structure $< \mathcal{W}_k, p_k, P, i_k, \mathcal{A} >$ is a generalized incidence calculus theory. In the case that the set of possible worlds is a joint space of several probabilistic assumption sets, labels of nodes need to be reconstructed. Following the above case if $S_{A_1,\ldots A_m} = \{A, \neg A\}$ and $S_{B_1,\ldots,B_n} = \{B, \neg B\}$, the label of node n_i can be changed as

$$\begin{aligned}
&L(n_i) = \{\{A\} \otimes \{B, \neg B\}, \{A, \neg A\} \otimes \{B\}\} \\
&= \{\{\{A,B\},\{A,\neg B\}\},\{\{A,B\},\{\neg A,B\}\}\} \\
&= \{\{A,B\},\{A,\neg B\},\{\neg A,B\}\}
\end{aligned}$$

In general, $L(n_i) = \{\{A\} \otimes S_{B_1,\ldots,B_n}, S_{A_1,\ldots,A_m} \otimes \{B\}\}$.

Step 5: for each premise node, create a generalized incidence calculus theory and add the set of possible worlds to the list. For example, for premise e, a suitable generalized incidence calculus theory might be $< \{V\}, \varrho(V) = 1, P, \{e\}, i_j(e) = \{V\} >$. The added probabilistic assumption set must be different from any set in the list.

Step 6: combining these generalized incidence calculus theories we have the result that for any derived node d_i, there is $i_*(d_i) \equiv FL(d_i) \setminus FL(\perp)$. $FL(d_i) \setminus FL(\perp)$ means deleting those conjunctive parts which appear in both $FL(d_i)$ and $FL(\perp)$.

So both the label set and the degree of belief in a node can be obtained in this combined generalized incidence calculus theory.

4.3 Formal proof

In this section we will give the formal proof about the equivalence between an ATMS and the transformed generalized incidence calculus theories.

Theorem 1 *Given an ATMS, there exists a set of generalized incidence calculus theories such that the reasoning result of the ATMS is equivalent to the result obtained from the combination of these theories. For any node d_l in an ATMS, $FL(d_l) \setminus FL(\perp)$ is equivalent to the lower bound of the incidence set of formula d_l in the combined generalized incidence calculus theory, that is $FL(d_l) \setminus FL(\perp) \equiv i_*(d_l)$. The nogood environments is equivalent to a subset of the set of possible worlds which causes conflicts, that is $FL(\perp) \equiv \mathcal{W}_0$.*

PROOF

The purpose of this proof is that, applying the Equivalent Transformation Algorithm in Definition 6 on a given ATMS, we get a list of generalized incidence calculus theories, the combined generalized incidence calculus theory of these theories generates the same label set and belief degree of a node as the ATMS does.

Assume that the nodes of an ATMS are divided into four sets, *e.g.*, a set of assumption nodes, a set of assumed nodes, a set of premise nodes and a set of derived nodes.

Step A: In order to carry out the proof below, we need to reconstruct the justifications of derived nodes to ensure that justifications of derived nodes contain only assumed nodes or premise nodes. This can be done as follows.

Given a derived node d_l, choose a node from its justifications. If the node is an assumption C, then create an assumed node c with single environment $\{C\}$ and single justification (C) and then replace C with c in any justifications where C appears. If the node is a derived node, then replace the node with the justifications of this node. For example if d_l is such a derived node with justifications $\{(z_1, z_2)(z_3, z_4)\}$ and d_l appears in a justification of node d_j as $\{(..., d_l, ...), ...\}$, then d_l is replaced with its justifications and the new justifications of d_j are $\{(..., z_1, z_2, ...), (..., z_3, z_4, ...), ...\}$.

Repeat this procedure until all nodes in the justifications of a derived node are either assumed nodes or premise nodes. As a consequence, an environment of a derived node contains only assumptions because labels of assumed and premise nodes contain only assumptions.

Step B: For any derived node d_l, suppose its justifications are

$$\{(a_1, a_2, ...), (b_1, b_2, ...), ...\}$$

then the conjunction of each justification of d_l implies d_l, such as $a_1 \wedge a_2 \wedge ... \rightarrow d_l$. If we denote this implication as $\models$, then we have $a_1 \wedge a_2 \wedge ... \models d_l$. If we let $j(d_l) = \{a_1 \wedge a_2 \wedge ..., b_1 \wedge b_2 \wedge ..., ...\}$ then $j(d_l) \models d_l$. The environments of d_l will be

$$(L(a_1) \otimes L(a_2) \otimes ...) \cup (L(b_1) \otimes L(b_2) \otimes ...) \cup ...$$

For example, if

$$L(a_1) = \{l_{i1}, l_{i2}, ...\}$$

and

$$L(a_2) = \{l_{j1}, l_{j2}, ...\}$$

then

$$L(a_1) \otimes L(a_2) = \cup_{t,k}\{l_{it} \cup l_{jk}\}$$

In general for a derived node d_l, assume that d_l has a justification $(n_1, n_2, ..., n_l)$, then

$$L(n_1) \otimes L(n_2) \otimes ... \otimes L(n_l) \setminus L(\perp)$$

is the label set of d_l.

Step C: After forming a language set from higher level nodes, a series of generalized incidence calculus theories (assume n theories in total) can be constructed from assumed nodes and premise nodes based on *steps 4 and 5* described in the Equivalent Transformation Algorithm. Any two sets of possible worlds of such theories are required to be probabilistically independent and all of them can be combined using Theorem 2 in Chapter 3 and the subset of possible worlds which leads to contradictions is W_0.

Suppose $(n_1, n_2, ..., n_l)$ is a justification of a derived node d_i (we have ensured that these nodes are either assumed nodes or premise nodes) and they are arranged into t generalized incidence calculus theories. Combining them we will obtain the generalized incidence calculus theory

$$< \mathcal{W}_1, \mu_1', P, \mathcal{A}_1', i_1' > \tag{3}$$

$$\begin{aligned}
i_1'(n_1 \wedge n_2 \wedge ... \wedge n_l) &= i_1(n_{11} \wedge ... \wedge n_{1m_1}) \otimes ... \otimes i_t(n_{t1} \wedge ... \wedge n_{tm_t}) \setminus W_1' \\
&= (L(n_{11}) \otimes ... \otimes L(n_{1m_1}) \otimes ... \otimes (L(n_{t1}) \otimes ... \otimes L(n_{tm_t})) \setminus W_1' \\
&= L(n_1) \otimes L(n_2) \otimes ... \otimes L(n_l) \setminus W_1'
\end{aligned}$$

where $\{n_1, ..., n_l\} = \{n_{11}, ..., n_{1m_1}, ...n_{t1}, ..., n_{tm_t}\}$ and $(n_{11} \wedge ... \wedge n_{1m_1})$, ..., $(n_{t1} \wedge ... \wedge n_{tm_t})$ are in these different generalized incidence calculus theories, and W_1' is the subset of possible worlds which leads to contradictions after combing these t generalized incidence calculus theories.

Assume that by combining the remaining $n - t$ generalized incidence calculus theories we have

$$< \mathcal{W}_2, \mu_2', P, \mathcal{A}_2', i_2' > \tag{4}$$

where $\mathcal{A}_2' = \{y_1, y_2, ..., y_n\}$ and the subset of possible worlds leading to contradictions is $\mathcal{W}_2'$. To combine the theories in (3) and (4), $\phi \wedge y_1, \phi \wedge y_2, ..., \phi \wedge y_n$ will be in the set of axioms of the new combined theory.

$$< \mathcal{W}_3, \mu_3', P, \mathcal{A}_3', i > \tag{5}$$

Here ϕ denotes $n_1 \wedge n_2 \wedge ... \wedge n_l$. Because $\phi \wedge y_j \models \phi$ and for any $\psi \wedge y_j \models \phi \wedge y_j$, $\psi \models \phi$, the following equation holds.

$$\begin{aligned}
i_*(\phi) &= \textstyle\bigcup_j i(\phi \wedge y_j) \\
&= \textstyle\bigcup_j i_1'(\phi) \otimes i_2'(y_j) \setminus W_3' \\
&= i_1'(\phi) \otimes \textstyle\bigcup_j i_2'(y_j) \setminus W_3' \\
&= i_1'(\phi) \otimes (\mathcal{W}_2 \setminus W_2') \setminus W_3' \qquad \text{as } \textstyle\bigcup_j i_2'(y_j) = \mathcal{W}_2 \setminus W_2' \\
&= i_1'(\phi) \otimes \mathcal{W}_2 \setminus ((i_1'(\phi) \otimes W_2') \cup W_3') \\
&= (L(n_1) \otimes L(n_2) \otimes ... \otimes L(n_l) \setminus W_1') \otimes \mathcal{W}_2 \setminus ((i_1'(\phi) \otimes W_2') \cup W_3') \\
&= (L(n_1) \otimes L(n_2) \otimes ... \otimes L(n_l)) \otimes \mathcal{W}_2 \setminus ((W_1' \otimes \mathcal{W}_2) \cup (i_1'(\phi) \otimes W_2') \cup W_3') \\
&= ((L(n_1) \otimes L(n_2) \otimes ... \otimes L(n_l)) \otimes \mathcal{W}_2) \setminus W_0
\end{aligned}$$

where W_3' is the set of possible worlds which leads to contradictions after combining the generalized incidence calculus theories i_1' and i_2'. The incidence function is i in the final generalized incidence calculus theory. W_0 is the total set of possible worlds causing conflict after combining all generalized incidence calculus theories.

Because of the relation $n_1 \wedge ... \wedge n_l \rightarrow d_l$ in the ATMS, we have the relation $n_1 \wedge ... \wedge n_l \rightarrow d_l$ in extended incidence calculus. So $i_*(\phi) \subseteq i_*(d_l)$. In general, if there are k justifications for node d_l, the environments obtained from k justifications are $(L(a_{11}) \otimes ... \otimes L(a_{1x})) \cup ... \cup (L(a_{k1} \otimes ... \otimes L(a_{ky})) \setminus L(\perp)$, then there are k corresponding formulae $\phi_1, \phi_2, ..., \phi_k$, where $i_{j*}(\phi_j) \subseteq i_*(d_l)$ for $j = 1, ..., k$. So $\bigcup_j i_{j*}(\phi_j) \subseteq i_*(d_l)$.

Step D: In the ATMS, a *nogood* environment is derived if $\perp$ is proved. When c and $\neg c$ are both derived, $L(c) \otimes L(\neg c)$ is a *nogood* environment. For any higher level node a, $(a, \neg a)$ is automatically recognized as a justification of node $\perp$ and $L(\perp) = nogood$. Certainly for an assumption A, $(A, \neg A)$ is also a justification of node $\perp$, but adding such justifications does not affect the result in our discussion, so in the following we only consider justifications of $\perp$ which are in the form of $(a, \neg a)$.

Choosing a justification of node $\perp$, such as $(c, \neg c)$, $L(c) \otimes L(\neg c)$ will be a part of environments of nogood. When c or $\neg c$ is a derived node, we replace c or $\neg c$ with its label. Suppose that the justifications of c are $\{(z_1, z_2, ...), (x_1, x_2, ...), ...\}$ and the justifications of $\neg c$ are $\{(y_1, y_2, ...), ...\}$, then $\{(z_1, z_2, ..., y_1, y_2, ...), (x_1, x_2, ..., y_1, y_2, ...), ...\}$ will be the justifications of $\perp$. Therefore $(L(z_1) \otimes L(z_2) \otimes ... \otimes L(y_1) \otimes L(y_2) \otimes ...) \cup (L(x_1) \otimes L(x_2) \otimes ... \otimes L(y_1) \otimes L(y_2) \otimes ...)$ are nogood environments. Because $z_1 \wedge z_2 \wedge ... \wedge y_1 \wedge ... =\perp$ and $x_1 \wedge x_2 \wedge ... \wedge y_1 \wedge ... =\perp$, we have $(L(z_1) \otimes L(z_2) \otimes ... \otimes L(y_1) \otimes ...) \cup (L(x_1) \otimes L(x_2) \otimes ... \otimes L(y_1) \otimes ...) \subseteq W_0$ based on **Step C** above. Therefore $FL(\perp) \subseteq W_0$.

The other way around, for any element $w \in W_0$, in the combined theory there exists a formula $\phi_1 \wedge \phi_2 \wedge ... \wedge \phi_n =\perp$ and $w \in L(\phi_1) \otimes ... \otimes L(\phi_n)$. Deleting those ϕ_j which will not destroy the equation $\wedge_i \phi_i =\perp$, we will have $\psi_1 \wedge ... \wedge \psi_m =\perp$. Therefore there exists a node z, the conjunction of some ψ_i implies z and the conjunctions of remaining ψ_j implies $\neg z$. So $z \wedge \neg z = \psi_1 \wedge ... \wedge \psi_m =\perp$, and $L(\psi_1) \otimes ... \otimes L(\psi_m)$ are nogood environments. It is straightforward that w is in the full extension of $L(\psi_1) \otimes ... \otimes L(\psi_m)$, so w is a nogood environment, that is $FL(\perp) \supseteq W_0$, so $FL(\perp) = W_0$.

Step E: Using the result from **Step C** and **Step D**, because $\bigcup_j i_{j*}(\phi_j) \subseteq i_*(d_l)$, we have the following non-equations.

$$((L(a_{11}) \otimes ... \otimes L(a_{1x})) \otimes ... \otimes (L(a_{k1} \otimes ... \otimes L(a_{ky}))) \setminus W_0 \subseteq i_*(d_l)$$

$$FL(d_l) \setminus FL(\perp) \subseteq i_*(d_l)$$

The other way around, for any $w \in i_*(d_l)$, there exists a formula $\phi = \phi_1 \wedge ... \wedge \phi_n$ and $w \in i(\phi)$. There is also a formula $\psi \in FL(d_l)$ such that $\psi = \psi_1 \wedge ... \wedge \psi_m$, $\phi \to \psi$. So $w \in i_*(\psi) = L(\psi_1) \otimes ... \otimes L(\psi_m) \setminus W_0$. Based on the definition of $FL(d_l)$, $\psi_1 \wedge ... \wedge \psi_m$ should be a justification of node d_l, so $L(\psi_1) \otimes ... \otimes L(\psi_m) \setminus L(\perp)$ will be the environments of d_l. Therefore w is in the full extension of $FL(d_l) \setminus FL(\perp)$. That is $FL(d_l) \setminus FL(\perp) \supseteq i_*(d_l)$, so eventually $FL(d_l) \setminus FL(\perp) = i_*(d_l)$.

QED

Example 6

Example 6 shows the way of dealing with conflict information. Following the story in Example 5, suppose we are told later that f is also observed and there is a rule $f \to \neg c$ with degree 0.8 in the knowledge base. That is, three more nodes in the ATMS are used.

assumed node: $< f \to \neg c, \{\{U\}\}, \{(U)\} >$
premise node: $< f, \{\{\}\}, \{()\} >$

assumption node: $< U, \{\{U\}\}, \{(U)\} >$
pas: $S_{U,\neg U} = \{U, \neg U\}$, $S_{F,\neg F} = \{F, \neg F\}$.

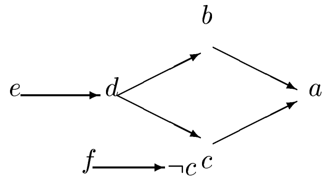

Figure 2. semantic network of inference rules

Here *pas* means probabilistic assumption set and $S_{F,\neg F}$ is created to support premise node f.

In the ATMS, we can infer that one environment of node c is $\{E, Z, Y\}$ and one environment of node $\neg c$ is $\{F, U\}$. So the *nogood* environment is $\{E, X, Y, F, U\}$. The belief in node a needs to be recalculated in order to re-distribute the weight of conflict on other nodes. The new belief in node a is 0.366 given in [16].

In extended incidence calculus similar to Example 5, two more generalized incidence calculus theories are constructed from the assumed node $f \to \neg c$ and the premise node f. Combining these two theories with the final one we obtained in Example 1, we have $W_0 = \{UZY\}$[8], $i_*(a) = \{ZXV \cup ZYW\} \setminus W_0$. Therefore $wp(\{UZY\}) = 0.48$ which is the weight of conflict and $p'_*(a) = wp(\{ZXV \cup ZYW\}) \setminus \{UZY\}) = 0.366$ which is our belief in a. Both of these results are the same as those given in [16], but the calculation of belief in node a and the weight of conflict are based on extended incidence calculus.

4.4 Comparison with Laskey and Lehner's work

The work carried out in this section has some similarity with Laskey and Lehner's work in [16]. The key idea in [16] is mainly about to create the medium level elements between a set of beliefs and numerical assignments and then associate the numerical assignments to the medium level elements. The medium level elements are exactly the set of possible worlds in extended incidence calculus and the set of assumptions in an ATMS. Both of our and Laskey and Lehner's work try to group assumptions into different sets and each set is associated with a probability distribution. Both of the work calculate labels and degrees of belief in nodes. They all concern the normalization after conflict is discovered and the total conflict weight is obtained. However the result we presented here is more theoretical. We provided a formal proof on the connections between extended incidence calculus and the ATMS while Laskey and Lehner didn't. Moreover, the result obtained in this section provides a theoretical basis for some results obtained in [16]. In this subsection, we will explain this point in more detail.

Difference 1). In [16] after the label of a node is obtained, in order to calculate the belief in this node, an algorithm is given to rewrite a label as a list of disjoint conjuncts of assumptions. For instance, in Example 5 the label of node a is rewritten as $L(a) = \beta_1 \vee \beta_2$ where $\beta_1 = W \wedge Y \wedge Z$ and $\beta_2 = (V \wedge X \wedge Z \wedge \neg W) \vee (V \wedge X \wedge Z \wedge W \wedge \neg Y)$.

If we simplify the elements in the full extension of a label (i.e. using Z to replace $(Z \wedge \neg W) \vee (Z \wedge W)$), we can get exactly those β list required in [16].

[8]In order to state the problem clearly, we use UZY instead of $UZYS_XS_WS_VS_ES_F$.

Difference 2). In [16] when *nogood* environments are produced, the beliefs in nodes are calculated in the following way

$$Bel(node) = \frac{Pr(label \cap \neg nogood)}{Pr(\neg nogood)} = \frac{Pr(label \cap \neg nogood)}{1 - Pr(nogood)}$$

It is suggested that the whole *nogood* environments can be divided into two groups $nogood_1$ and $nogood_2$ where $nogood_2$ has no overlap with environments in $nogood_1$ or $label$. So in a real calculation *nogood* is replaced by $nogood_1$ and it is claimed that such replacement doesn't affect the whole result. They didn't provide a proof. We will prove this result is sound.

Theorem 2 *Assume that all nogood environments can be divided into two disjoint groups* $nogood_1$ *and* $nogood_2$. *For a node* d_l, *if* $L(d_l)$ *has no overlap with* $nogood_2$, *then the following equation holds.*

$$Bel(d_l) = \frac{Pr(L(d_l) \cap nogood)}{1 - Pr(nogood)} = \frac{Pr(L(d_l) \cap nogood_1)}{1 - Pr(nogood_1)}$$

PROOF

If all *nogood* environments can be divided into two disjoint groups, then it is possible to divide all the corresponding generalized incidence calculus theories into two groups based on **Step C** in section 4.3. The combination of generalized incidence calculus theories in two groups produces two conflict sets, referred to as $nogood_1$ and $nogood_2$ respectively. The final combination of these two generalized incidence calculus theories will not produce any conflict sets (if it does then the assumption that $nogood_1$ and $nogood_2$ are disjoint is wrong). Assume that the two generalized incidence calculus theories are i_1 and i_2 respectively after combining two groups of generalized incidence calculus theories, for a formula ϕ, if the list of axioms making ϕ true are $x_1, x_2, ..., x_n$, then

$$i_*(\phi) = \bigcup_j (i_1(x_j))$$

Assume that the list of all axioms for incidence function i_2 are $y_1, y_2, ..., y_m$, then combining i_1 and i_2 we have

$$\begin{aligned}
i'_*(\phi) &= \cup_l(\cup_j i(x_l \wedge y_j)) \\
&= \cup_l(\cup_j i_1(x_l) \otimes i_2(y_j)) \\
&= \cup_l(i_1(x_l) \otimes \cup_j i_2(y_j)) \\
&= \cup_l(i_1(x_l) \otimes (\mathcal{W}_2 \setminus FL(nogood_2))) \\
&= (\cup_l i_1(x_l)) \otimes (\mathcal{W}_2 \setminus FL(nogood_2)) \\
&= i_*(\phi) \otimes (\mathcal{W}_2 \setminus FL(nogood_2))
\end{aligned}$$

So $p_*(\phi) = \mu(i'_*(\phi)) = \mu(i_*(\phi)) \times \mu(\mathcal{W}_2 \setminus FL(nogood_2)) = \mu(i_*(\phi))$. That is

$$Bel(\phi) = \frac{Pr(L(\phi) \cap nogood_1)}{1 - Pr(nogood_1)}$$

Therefore, those nogood environments which don't have overlap with the label of a node don't affect the belief in this node.

END

Difference 3). The major step in [16] is to create an auxiliary set for each belief function and let the auxiliary set carry the information provided by the belief function. So the probability distribution on an auxiliary set which in turn gives the belief function on another set can be thought as the source for this belief function. Therefore the two auxiliary sets defined in this way should be DS-Independent, otherwise these two belief functions cannot be combined by the Dempster's Rule and the result obtained in an ATMS has no way to compare with the result in DS theory.

However, in extended incidence calculus, we don't need to make such an assumption. For dependent probabilistic assumption sets, as long as we can find their joint probabilistic assumption set, we can still combine them using the rule in [17], [18]. If there are a number of probabilistic assumption sets and some of them are dependent, we combine dependent probabilistic assumption sets first and then carry out the combination for the rest.

Example 7

Example 7 demonstrates the point we discussed in 2) above. Assume that the ATMS network is extended as in Figure 4 by adding more nodes in it. When the facts h and j are observed, both i and $\neg i$ will be derived, then there will be a conflict. So the total *nogood* environments are $\{UZY, HI\}$. Without giving any obvious links between $h \rightarrow i$, $j \rightarrow \neg i$ and the previous network, $\{HI\}$ should have no effect on the belief in a. So the belief in a shouldn't be changed even more facts are observed.

assumed nodes: $< h \rightarrow i, \{\{H\}\}, \{(H)\} >$
$< j \rightarrow \neg i, \{\{I\}\}, \{(I)\} >$
premise nodes: $< h, \{\{\}\}, \{()\} >$
$< j, \{\{\}\}, \{()\} >$
assumption node: $< H, \{\{H\}\}, \{(H)\} >$
$< I, \{\{I\}\}, \{(I)\} >$
pas: $S_{H,\neg H} = \{H, \neg H\}$, $S_{I,\neg I} = \{I, \neg I\}$.
$S_{G,\neg G} = \{G, \neg G\}$, $S_{L,\neg L} = \{L, \neg L\}$.

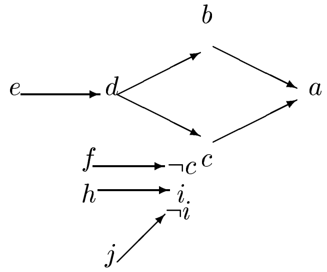

Figure 4. Extending the existing ATMS

If we wish to consider this problem in extended incidence calculus, after we encoded the new assumed and premise nodes into incidence calculus theories, the combination of these theories produces a conflict set $W'_0 = \{HI\}$. The further combination of this theory with the generalized incidence calculus theory obtained in Example 5 gives the final result of the impact of all evidence. In this final generalized incidence calculus theory, we have $p''_*(a) = p'_*(a) = 0.366$ while the whole weight of conflict is

$wp(FL(UZY \cup HI))$
$p_U(U)p_Z(Z)p_Y(Y) + p_H(H)p_I(I) - p_U(U)p_Z(Z)p_Y(Y)p_H(H)p_I(I)$
$= 0.48 + p_H(H)p_I(I) - 0.48p_H(H)p_I(I)$

Therefore in extended incidence calculus we don't need to divide *nogood* environments into different groups while the correct result can still be achieved.

5 EXTENDED INCIDENCE CALCULUS CAN PROVIDE JUSTIFICATIONS FOR THE ATMS

In the previous sections, we have discussed the formal relations between extended incidence calculus and the ATMS. The major similarity of the two reasoning mechanisms is that the justifications in an ATMS are equivalent to the essential semantic implication sets in incidence calculus. As a result, the labels of nodes are equivalent to the incidence sets of the corresponding nodes. However, a difference between these two reasoning patterns is that the justifications are assigned by the designers in an ATMS while essential semantic implication sets are discovered automatically in extended incidence calculus. Therefore, the whole reasoning procedure in extended incidence calculus is automatic while the one in an ATMS is semi-automatic. The procedure of discovering semantic implication sets in extended incidence calculus can be regarded as a tool to provide justifications for an ATMS. The application of this procedure into an ATMS can release a system designer from the task of assigning justifications and this procedure can guarantee those justifications are non-redundant. A problem with this procedure is that it is slow to find all essential semantic implication sets. If it is possible to have a fast algorithm for this procedure, then an ATMS can be established and extended automatically without a designer's involvement.

We use an example to show our idea here concretely.

Example 10 Providing justifications automatically using extended incidence calculus

We examine Example 5 in [16] in a different way here. Assume that our objective in Example 5 is to calculate the impact on a when e is observed. Because there is no direct effect from e on a, a diagram shown as Figure 1 is created to build a link between e and a. In order to infer a, the justifications for node $e \to a$ are essential to be given in an ATMS. Assume that the information carried by this diagram is denoted as S_I and the information specifying justifications is denoted as S_J, then in an ATMS we have

$$S_I \cup S_J \Rightarrow L(e \to a) \tag{6}$$

Here notation $A \Rightarrow B$ means that from information carried by A, it is possible to infer information carried by B through some logical methods. S_J may either contain the justifications for node $e \to a$ only or consists of more justifications for the assisting nodes (such as $e \to b$). We say that S_J is the extra information for the system inference.

Given the same initial information carried by S_I to it, extended incidence calculus does inferences without requiring any more information. The inference procedure produces

$$S_I \Rightarrow i_*(e \to a) \cup ESI(e \to a)$$

This can be explained as from information in S_I, we can obtain both the lower bound of the incidence set and the inference pathes of a node. The essential semantic implication set for a node contains exactly the justifications for the same node. Therefore the extra information required by the ATMS can be supplied by extended incidence calculus as an output in general and we are able to change (6) as follows in an ATMS

$$S_I \cup ESI(e \to a) \Rightarrow L(e \to a)$$

which takes the output from extended incidence calculus as an input in the ATMS.

So we can abstract out essential semantic implication sets for all necessary formulae and assign them on the corresponding nodes without considering assumptions on the initial nodes. In this way, an justification existing ATMS can be constructed.

So we conclude that the inference result in extended incidence calculus provides justifications for an ATMS automatically.

6 CONCLUSION

A notable statement about the relations between the ATMS and extended incidence calculus has been given by Pearl [20]. He said:"In the original presentation of incidence calculus, propositions were not assigned numerical degrees of belief but instead were given a list of labels called *incidences*, representing a set of situations in which the propositions are true. ... Thus, incidences are semantically equivalent to the ATMS notion of 'environments', and it is in this symbolic form that incidence calculus was first implemented by Bundy." In this paper we have discussed the relations intensively. This discussion proves the equivalence between extended incidence calculus and the ATMS. The result tells us that extended incidence calculus itself is a unification of both symbolic and numerical approaches. It can therefore be regarded as a bridge between the two reasoning patterns. This result also gives theoretical support for research on the unification of the ATMS with numerical approaches. In extended incidence calculus structure, both symbolic supporting relations among statements and numerical calculation of degrees of belief in different statements are explicitly described. For a specific problem, extended incidence calculus can either be used as a support based symbolic reasoning system or be applied to deal with numerical uncertainties. This feature cannot be provided by pure symbolic or numerical approaches independently.

An advantage of using extended incidence calculus to make inferences is that it doesn't require the problem solver to provide justifications. The whole reasoning procedure is performed automatically. The inference result can be used to produce the ATMS related justifications. The calculation of degrees of beliefs in nodes is based on the probability distributions on assumption sets which can either be dependent or independent.

In the traditional TMS or ATMS, when nogood environments are generated, a number of assumptions need to be deleted (or the truth value of the assumptions are changed to be false) in order to restore the consistency in the whole system. This procedure is usually called belief revision [8], [11], [12]. Notions of epistemic entrenchment are used to determine which sets of assumptions to favour over others when resolving a conflict. It should be interesting to use the extended incidence calculus as a means of supplying a formal basis for this principle.

Acknowledgement

The first author is being supported by a "Colin and Ethel Gordon Scholarship" of Edinburgh University and an ORS award.

References

[1] Bundy,A., Incidence calculus: A mechanism for probabilistic reasoning. *Journal of Automated Reasoning* 1:263-83, 1985.

[2] Bundy,A., Incidence calculus, *The Encyclopedia of AI.* 663-668, 1992.

[3] d'Ambrosio,B., A hybrid approach to reasoning under uncertainty, *Int. J. Approx. Reasoning* 2 (1988): 29-45.

[4] d'Ambrosio,B., Incremental Construction and Evaluation of Defensible Probabilistic Models, *I.J.Approx. Reasoning* 4 (1990): 233-260.

[5] Dempster,A.P., Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Stat..* 38, 325-339.

[6] de Kleer,J., An assumption-based TMS. *Artificial Intelligence* 28 (1986) 127-162.

[7] de Kleer,J. and B.C.Williams, Diagnosing multiple faults, *Artificial Intelligence* 32 (1987) 97-130.

[8] Doyle,J., A truth maintenance system. *Artificial Intelligence* 12 (3): 231-72, 1979.

[9] Doyle,J., Reason maintenance and belief revision: foundation vs. coherence theories., *Belief Revision*, (P.Gardenfors Ed.) Cambridge University Press, 29-51, 1992.

[10] Dubois,D., J.Lang and H.Prade, Handling uncertain knowledge in an ATMS using possibilistic logic, *ECAI-90 workshop on Truth Maintenance Systems*, (1990) Stockholm, Sweden.

[11] Dubois,d. and H.Prade, Belief change and possibility theory, *Belief Revision*, (P.Gardenfors Ed.) Cambridge University Press, 142-182, 1992.

[12] Gardenfors,P., Belief revision: an introduction, *Belief Revision*, (P.Gardenfors Ed.) Cambridge University Press, 1-28, 1992.

[13] Falkenhainer, B., Towards a general purpose belief maintenance system, in: Uncertainty in Artificial Intelligence, 2 J.F.Lemmer, L.N.Kanal, Eds. North-Holland, Amsterdam, 1988.

[14] Fringelli, B., Marcugini, S., Milani, A., Rivoira, S. A reason maintenance system dealing with vague data. *Proc. of 7th Conference on Uncertainty in Artificial Intelligence* D.D'Ambrosio, Ph.Smets, P.Bonissone, Eds. pp111-117, 1991.

[15] Fulvio Monai,F. and T.Chehire, Possibilistic Assumption based Truth Maintenance Systems, Validation in a Data Fusion Application, *Proc. of the eighth conference on uncertainty in artificial intelligence.* Stanford, 83-91, 1992.

[16] Laskey,K.B. and P.E.Lehner, Assumptions, Beliefs and Probabilities, *Artificial Intelligence* 41 (1989/90) 65-77.

[17] Liu,W., *Extended incidence calculus and its comparison with related theories*, unpublished PhD thesis. Dept. of AI, University of Edinburgh, 1995.

[18] Liu,W. and A.Bundy, A comprehensive comparison between generalized incidence calculus and the Dempster-Shafer theory of evidence, *I.J.of Human-Computer Studies* (1994) **40**:1009-1032.

[19] Liu,W., A.Bundy and Dave Robertson, Recovering incidence functions. In the *Procs. of Symbolic and quantitative Approaches to Reasoning and Uncertainty.* Lecture notes in computer science 747, pp249-256, Springer, 1993.

[20] Pearl,J., *Probabilistic Reasoning in Intelligence Systems: networks of plausible inference.* Morgan Kaufmann Publishers, Inc., 1988.

[21] Proven,G.M., An analysis of ATMS-based techniques for computing Dempster-Shafer belief functions. *Proc. of the 11th International Joint Conf. on Artificial Intelligence*, p:1115-1120, 1989.

[22] Shafer,G., *A mathematical theory of evidence* Princeton University Press, 1976.

[23] Voorbraak,F., On the justification of Dempster's rule of combination. *Artificial Intelligence.* 48, 171-197, 1991.