The Incidence Propagation Method

W. Liu
School of Information and Software Engineering
University of Ulster
Co.Antrim BT37 0QB, UK

Abstract

Incidence calculus is a probabilistic logic developed from propositional logic which associates probabilities with sets of possible worlds rather than with formulae directly. The probability of a formula is defined as the probability of the set of possible worlds in which this formula is true. This set of possible worlds is named as the incidence set of this formula. So the task of calculating probabilities of formulae relies on the task of obtaining incidence sets for formulae. In this paper, we present an approach for manipulating incidences in *extended incidence calculus* in the situation that the language set is large. We will show how to decompose this large set into small, but coherent sets and then how to propagate incidences among these sets. In this way incidence sets can be calculated efficiently.

1 Introduction

1.1 Background

Unlike other pure numerical uncertainty reasoning techniques, in incidence calculus [1], [2], numerical uncertainty values are associated with formulae indirectly through a set of possible worlds. Incidences of a formula are those possible worlds in which the formula is true. Incidences make incidence calculus possess rich properties whether incidence calculus is treated as a numerical uncertainty reasoning mechanism [4] or as a symbolic reasoning mechanism [5], [6]. The incidence plays an important role in linking these two reasoning systems together [7]. So how to propagate incidences efficiently within a propositional language is crucial in applying incidence calculus.

Suppose that P is a set of propositions from which a set of formulae is derived using logical connectors and the set is denoted as $\mathcal{L}(P)$. Let $\mathcal{W}$ be a set of possible worlds, a subset W_j of $\mathcal{W}$ is said to be the incidence set of formula ϕ if each possible world $w \in W_j$ supports the truth of ϕ and W_j is the largest of this kind.

If we use $p(\phi)$ to denote the probability of ϕ, then $p(\phi)$ is defined as

$$p(\phi) = \mu(W_j) = \Sigma_{w \in W_j} \mu(w)$$

where μ is the probability distribution on $\mathcal{W}$ and for each w in $\mathcal{W}$, $\mu(w)$ is known.

So in order to obtain the probability of a formula, we have to get the incidence set of this formula first. In incidence calculus, some initial relations between formulae and possible worlds are specified using an incidence function i. $i(\phi) = W_j$ means that the incidence set of formula ϕ is W_j. Usually, i is a partial function between $\mathcal{L}(P)$ and $\mathcal{W}$. That is, i is only defined on a subset $\mathcal{A}$ of $\mathcal{L}(P)$. $\mathcal{A}$ is called the set of *axioms*.

For Example, if it is assumed that $\mathcal{A} = \{A, B\}$ and the incidence sets of A and B are W_1 and W_2 respectively, then the incidence sets of $A \wedge B$ and $A \vee B$ are $W_1 \cap W_2$ and $W_1 \cup W_2$ respectively. This feature is called *truth functional* in incidence calculus. That is, the incidence set of a formula ψ ($\psi \in \mathcal{L}(A)$) is derivable. However, for any other formulae in $\mathcal{L}(P) \setminus \mathcal{L}(A)$, it is only possible to calculate the lower and upper bounds of the incidence sets of these formulae. In [1], a technique called the *Legal Assignment Finder* was designed in which a set of inference rules were defined to calculate the bounds based on an initial incidence assignment. It was proved that the exhaustive application of these rules will terminate.

1.2 Problems with incidence propagation

In [7], the original incidence calculus is extended in three aspects, *i.e.*, the representation of a wider range of information by requiring fewer conditions using generalized incidence calculus theories; a new algorithm for incidence assignments based on the numerical assignment; the combination rule to combine multiple pieces of evidence. The advanced reasoning systems is called *extended incidence calculus*. Extended incidence calculus has the advantage of being able to represent ignorance [4] but loses the feature of truth functionality. So the set of inference rules for deriving bounds of incidences in the *Legal Assignment Finder* in the original incidence calculus cannot be applied any more. The only method in extended incidence calculus for calculating bounds of incidences is based on logical entailment among formulae.

For instance, assume that the formulae in the set of axioms $\mathcal{A}$ are assigned incidences initially, *i.e.*, $i(\phi)$ is known for $\phi \in \mathcal{A}$, then for any formula ϕ, the lower bound of ϕ, denoted as $i_*(\phi)$, will be obtained based on the following equation:

$$i_*(\phi) = \cup_{\psi \models \phi, \psi \in \mathcal{A}} i(\psi) \qquad (1)$$

where $\psi \models \phi$ means that $\psi \rightarrow \phi$ is a tautology for $\psi \in \mathcal{A}$.

If $\mathcal{A}$ is large, this method would be very inefficient. This is the first problem that arises in extended incidence calculus.

Another problem in either the original or extended incidence calculus is that there is only one set of propositions constructed and manipulated for a given problem. Therefore any relevant statements (propositions) will have to be included in it.

Example 1.1

Assume that we have the following sets of propositions which are associated with three different questions:

$$\begin{pmatrix} \textit{Sunny} \\ \textit{Windy} \\ \textit{Rainy} \end{pmatrix} \Longrightarrow \begin{pmatrix} \textit{Have a picnic} \\ \textit{Go to work} \end{pmatrix} \Longrightarrow \begin{pmatrix} \textit{Wear a suit} \\ \textit{Casual cloths} \end{pmatrix}$$

The elements in the first set are mapped to some elements in the second set, and the elements in the second set are mapped to the elements in the third set. Notation $\Longrightarrow$ between two sets means that some kind of mapping relations is established among the elements in the two sets. We name these three sets as P_1, P_2 and P_3 respectively.

Suppose that we also have the following statements among the propositions in the three sets:

S_1: If it is a sunny day, Then I will go out to have a picnic;

S_2: If it is rainy or windy, Then I will go to work;

S_3: If I go to work, Then I will have to wear a suit.

In either the original or extended incidence calculus, a set of propositions, P, is the union of P_1, P_2, P_3, $P = \{$Sunny, Rainy, Windy, Have a picnic, Go to work, Wear a suit$\}$ and the corresponding $\mathcal{L}(P)$ can then be formed which contains these propositions and any other formulae produced from them using logical connectors.

Assume that we observe the fact 'it is rainy', then we can infer 'Go to work' based on sets P_1, P_2 and S_2 without involving the third set P_3. If we observe the fact that 'Go to work', then we can infer 'Wear a suit' based on sets P_2, P_3 and S_3 without involving set P_1. This tells us that for some cases only part of $\mathcal{L}(P)$ is useful for the inference. Manipulating this part of relevant information rather than the whole set would certainly increase the efficiency of inference.

Reconsidering these three sets, we conclude that if we keep these three sets separately, then if a piece of evidence is first put on P_1, it can be propagated to P_2 when necessary and further to P_3 if required through some mapping relations among these sets. This would be particularly useful when P is considerably large.

Based on this analysis, it suggests that an appropriate approach for propagating incidences from one set to another is necessary in order to reduce any worthless inferences. However, this suggestion brings up a number of questions:

How to form different sets for a particular problem?

How to propagate incidences from one set to another?

Is the propagated message still in the form of generalized incidence calculus theories if the original information is?

These are the main concerns of this paper. We propose that incidences can be propagated among different $\mathcal{L}(P)$s through some proper mapping relations. In this way, incidences can not only be calculated for formulae within one $\mathcal{L}(P)$ set but also be propagated to other different $\mathcal{L}(P)$ sets.

The paper is organized as follows. Section 2 introduces the basics of extended incidence calculus. Section 3 structures the problem with incidence propagation. Section 4 discusses the methods for efficient propagation of incidences. Section 5 analyses the approach proposed in the paper. Finally, in section 6, we summarize the paper.

2 Basics of Extended Incidence Calculus

We will refer to the following relevant definitions in the rest of the paper. More details can be found in [7].

Definition 1 *Propositional Language*

Let $P = \{q_1, q_2, ..., q_n\}$ be a finite set of propositions:

- *A set $\mathcal{A}t = \{\delta_j = q'_1 \wedge q'_2 \wedge ... \wedge q'_n\}$ is called the basic element set of P where q'_l is either q_l or $\neg q_l$.*
- *$\mathcal{L}(P)$ is called the language set of P by applying logical connectors $\wedge, \vee, \rightarrow$ and $\neg$ on propositions in P. Each formula ϕ in $\mathcal{L}(P)$ can be equivalently rewritten as the disjunction of elements in $\mathcal{A}t$ and $\mathcal{L}(P)$ is finite if the formulae are written in this way.*

Definition 2 *Generalized Incidence Calculus Theories*

A quintuple $< \mathcal{W}, \mu, P, \mathcal{A}, i >$ is called a generalized incidence calculus theory where

- *$\mathcal{W}$ is a finite set of possible worlds and μ is the discrete probability distribution on $\mathcal{W}$.*
- *$\mathcal{A}$ is a subset of $\mathcal{L}(P)$ which is closed under operator $\wedge$.*
- *i is a mapping function from $\mathcal{A}$ to $\mathcal{W}$ which must satisfy the following three conditions:*

$$i(true) = \mathcal{W}$$
$$i(false) = \{\}$$
$$i(\phi \wedge \psi) = i(\phi) \cap i(\psi)$$

It is assumed that *true* is an axiom in any generalized incidence calculus theory and included in $\mathcal{A}$. In the case that a formula ϕ is a tautology, ϕ should be an axiom and be included in $\mathcal{A}$ as well.

Definition 3 *Lower and Upper Bounds of Incidences*

For a formula $\varphi \in \mathcal{L}(P) \setminus \mathcal{A}$, it is only possible to obtain the lower and upper bounds for it using the following equations:

$$i_*(\varphi) = \cup_{\phi \models \varphi, \phi \in \mathcal{A}} i(\phi)$$
$$i^*(\varphi) = \cap_{\varphi \models \phi, \phi \in \mathcal{A}} i(\phi)$$

It is easy to see that for any $\psi \in \mathcal{A}$, $i_*(\psi) = i(\psi)$.

Definition 4 *Lower and Upper Bounds of Probabilities*

For a formula $\varphi \in \mathcal{L}(P)$, we define the lower and upper bounds of its probability as

$$p_*(\varphi) = \mu(i_*(\varphi))$$
$$p^*(\varphi) = \mu(i^*(\varphi))$$

When $p_*(\varphi) = p^*(\varphi)$ for each $\varphi \in \mathcal{A}$, we define $p(\varphi) = p_*(\varphi)$ and $p(\varphi)$ is called the probability of φ.

To see how to make inference in extended incidence calculus, we look at an example.

Example 2.1 (from [3])

> A person has four coats: two are blue with single-breasted, one is grey and double-breasted and one is grey and single-breasted. To choose which colour of coat to wear, this person tosses a (fair) coin. Once the colour is chosen, which specific coat is wore is determined by a mysterious procedure. What is the probability of the person wearing a single-breasted coat?

To solve this problem in extended incidence calculus, we need to construct a generalized incidence calculus theory first. We let a set of propositions P be $P = \{grey, double\}$ where *grey* stands for 'The coat is grey' and *double* stands for 'The coat is double-breasted' and let $\mathcal{W} = \{w_1, w_2\}$ where w_1 for blue coats and w_2 for grey coats. Then we have

$\mathcal{A}t = \{grey \wedge double, \neg grey \wedge double, grey \wedge \neg double, \neg grey \wedge \neg double\}$

It is possible to derive that w_1 supports formula $\neg grey \wedge \neg double$ and w_2 supports formula $(grey \wedge \neg double) \vee (grey \wedge double)$. Therefore, we get a generalized incidence calculus theory as $< \mathcal{W}, \mu, P, \mathcal{A}, i >$ where

$\mu(w_1) = \mu(w_2) = 0.5$

$\mathcal{A} = \{\neg grey \wedge \neg double, (grey \wedge \neg double) \vee (grey \wedge double)\}$

$i(\neg grey \wedge \neg double) = \{w_1\}$

$i((grey \wedge \neg double) \vee (grey \wedge double)) = \{w_2\}$

So

$$i_*(\neg double) = i(\neg grey \wedge \neg double)$$
$$i^*(\neg double) = \mathcal{W} \setminus i_*(double) = \mathcal{W} \text{ and}$$
$$p_*(\neg double) = 0.5 \qquad p_*(\neg double) = 1$$

The answer to the question is that the probability of the person wearing a single-breasted coat lies between 0.5 and 1.

3 Structuring the Problem

In a real world problem specification, propositions may be divided into different sets rather than in one set. For instance, when we are concerned with a ship's location, we can form a set of propositions regarding its locations. When we are concerned with the ship's activities, we can form another set of propositions for its activities [8]. These two sets may or may not be used at the same time, so putting them together to form a large set causes great computational problem in most of the time. It is more reasonable to keep these two sets separately and then propagate incidences from one set to another. For example, we know that a ship's activity very much related to its location. When we know the incidences of its locations, we could get the incidences of its activities through relations on what location related to which activity.

If we form a single set P, difficulties arise when P is large. Assume that a set P consists of n propositions, let $m = 2^n$, then $\mathcal{L}(P)$ contains

$$1 + \frac{m}{1!} + \frac{m(m-1)}{2!} + ... + \frac{m!}{m!} = 2^m = 2^{2^n}$$

formulae.

If a set of axioms $\mathcal{A}$ has l elements, then it involves l steps of implication checking for $\varphi \in \mathcal{L}(P) \setminus \mathcal{A}$ in order to obtain its bounds of incidences. However, if there are t tautologies (in set S) in $\mathcal{L}(P)$, then the final set of axioms would be $(\mathcal{A} \cup S)^\wedge$ which is closed under $\wedge$. When $\mid (\mathcal{A} \cup S)^\wedge \mid$ is large, the computational complexity is certainly a big problem for the system inference. If it is possible to split P into small, but self-contained sets, among which mapping relations can be established, the inference procedure could be more efficient.

In summary, the problem with incidence propagation can be regarded as the problem of incidence propagation between sets. The problem then can be stated formally as follows.

Statement of the Problem. We are given two (or more) sets of propositions, the elements of which have some mapping relations and we know the incidences of some formulae in one set. We wish to calculate the incidences (or bounds) of formulae in another set through the mapping relations among elements provided.

4 The Incidence Propagation Method

Definition 5 *Let P_1 and P_2 be two sets of propositions, $\mathcal{A}t_1$ and $\mathcal{A}t_2$ be the two basic element sets. A function Γ is called a mapping function between $\mathcal{A}t_1$ and $\mathcal{A}t_2$ if for each element δ_j in $\mathcal{A}t_1$, there is a subset B of $\mathcal{A}t_2$ such that when δ_j is true then formula $\psi = \vee\beta_k$ ($\beta_k \in B$) is valid and B is the smallest set of this kind. We denote the mapping relation as $\Gamma : \mathcal{A}t_1 \to \mathcal{A}t_2$ and $\Gamma(\delta_j) = B$. Further, for a formula $\phi = \vee\delta_j$ where $\phi \in \mathcal{L}(P)$, we define $\Gamma(\phi)$ as $\Gamma(\phi) = \cup(B_j \mid \Gamma(\delta_j) = B_j)$.*

To simplify the statement, we use B_ψ to denote a subset B of $\mathcal{A}t_2$ when $\psi \equiv \vee\beta_k$ where $\beta_k \in B$. It is possible that different δ maps into the same subset B, but none of the elements is allowed to map to the empty set.

Usually, we require that P_1 and P_2 have no overlap (no common propositions), so $\mathcal{A}t_1$ and $\mathcal{A}t_2$ have no overlap as well. For instance, we can form a set of propositions, P_1, about the weather this week, we can also form a set of propositions, P_2, on going out for a picnic party this week. P_1 and P_3 have no overlap but can be linked together by a mapping relation if we want to choose a good day to be out.

Proposition 1 *Let ϕ_1 and ϕ_2 be two formulae in $\mathcal{L}(P_1)$, given a mapping relation between $\mathcal{A}t_1$ and $\mathcal{A}t_2$, we have*

$$\Gamma(\phi_1 \wedge \phi_2) \subseteq \Gamma(\phi_1) \cap \Gamma(\phi_2)$$

This Proposition can be proved as follows.

PROOF

Assume that $\phi_1 \wedge \phi_2$ is valid, then there exist ϕ'_1 and ϕ'_2 that $\phi_1 = \phi'_1 \vee (\phi_1 \wedge \phi_2)$ and $\phi_2 = \phi'_2 \vee (\phi_1 \wedge \phi_2)$ hold. So we have

$\Gamma(\phi_1) = \Gamma(\phi'_1) \cup \Gamma(\phi_1 \wedge \phi_2)$

$\Gamma(\phi_2) = \Gamma(\phi'_2) \cup \Gamma(\phi_1 \wedge \phi_2)$

and

$\Gamma(\phi_1) \cap \Gamma(\phi_2) = (\Gamma(\phi'_1) \cap \Gamma(\phi'_2)) \cup \Gamma(\phi_1 \wedge \phi_2)$

Therefore, $\Gamma(\phi_1 \wedge \phi_2) \subseteq \Gamma(\phi_1) \cap \Gamma(\phi_2)$.

When $\phi_1 \wedge \phi_2$ is false, the above relation still holds.

QED

Based on this proposition, we have the following theorem.

Theorem 1 *Let P_1 and P_2 be two sets of propositions and Γ be a mapping relation between the two corresponding basic element sets $\mathcal{A}t_1$ and $\mathcal{A}t_2$. Given an incidence function i_1 on a set of axioms $\mathcal{A}_1 \subseteq \mathcal{L}(P)$, a set of axioms $\mathcal{A}_2$ can be constructed as*

$\mathcal{A}_2 = \{\psi \mid \phi \in \mathcal{A}_1, \Gamma(\phi) = B_\psi\}$

A function i_2 is then defined on $\mathcal{A}_2$ as

$i_2(\psi) = \cup_{\Gamma(\phi)=B_\psi} i_1(\phi)$

Then i_2 is an incidence function on $\mathcal{A}_2$ after we add that $i_2(true) = \mathcal{W}$ and $i_2(false) = \{\}$.

PROOF To prove that i_2 is an incidence function, we only need to prove that i_2 is closed under $\wedge$, that is $i_2(\psi_1) \cap i_2(\psi_2) = i_2(\psi_1 \wedge \psi_2)$.

Assume that ψ_1, ψ_2 and $\psi_1 \wedge \psi_2$ are all in $\mathcal{A}_2$.

For $w \in i_2(\psi_1) \cap i_2(\psi_2)$, we have

$w \in i_2(\psi_1) \cap i_2(\psi_2) \Rightarrow$
$(\exists \phi_1, w \in i_1(\phi_1), \Gamma(\phi_1) = B_{\psi_1})$ and
$(\exists \phi_2, w \in i_1(\phi_2), \Gamma(\phi_2) = B_{\psi_2}) \Rightarrow$
$(\exists \phi_1, \phi_2) w \in i_1(\phi_1) \cap i_1(\phi_2) \Rightarrow$
$(\exists \phi_1, \phi_2) w \in i_1(\phi_1 \wedge \phi_2) \Rightarrow$
$(\exists \psi) w \in i_1(\phi_1 \wedge \phi_2)$,
$\Gamma(\phi_1 \wedge \phi_2) = B_\psi, \psi \models \psi_1, \psi \models \psi_2 \Rightarrow$
$(\exists \psi) w \in i_2(\psi), \psi \models \psi_1 \wedge \psi_2 \Rightarrow$
$(\exists \psi) w \in i_2(\psi_1 \wedge \psi_2)$

So $i_2(\psi_1 \wedge \psi_2) \subseteq i_2(\psi_1) \cap i_2(\psi_2)$.

On the other hand, from $w \in i_2(\psi_1 \wedge \psi_2)$, we have

$w \in i_2(\psi_1 \wedge \psi_2) \Rightarrow$
$(\exists \phi)\Gamma(\phi) = B_{\psi_1 \wedge \psi_2} = B_\psi, w \in i_1(\phi) \Rightarrow$
$(\exists \phi)\Gamma(\phi) = B_\psi \subseteq B_{\psi_1}$,
$B_\psi \subseteq B_{\psi_2}, w \in i_1(\phi), w \in i_2(\psi) \Rightarrow$
$w \in i_2(\psi_1), w \in i_2(\psi_2) \Rightarrow$
$w \in i_2(\psi_1) \cap i_2(\psi_2)$

So $i_2(\psi_1 \wedge \psi_2) \supseteq i_2(\psi_1) \cap i_2(\psi_2)$. Therefore, it is straightforward to say that $i_2(\psi_1 \wedge \psi_2) = i_2(\psi_1) \cap i_2(\psi_2)$.

QED

This theorem tells us that, when we know a piece of evidence about one language set in the form of a generalized incidence calculus theory, we can propagate this evidence onto another set through the mapping relations to form another generalized incidence calculus theory.

5 Analysing the approach

There are the following problems arising along with the establishment of the approach.

- Under which circumstances is the efficiency improved and what kind of worthless inferences are discarded.
- How meaningful is it to spit a whole set of propositions into a number of small sets and is it always possible?

To answer these questions, we look at the example in Section 1.2 again.

Example 5.1

Assume that there are three sets P_1, P_2 and P_3 as defined in Example 1.2. A set P is the union of P_1, P_2 and P_3. Given a set of possible worlds $\mathcal{W}$, suppose that the statement 'It is rainy' is supported by a subset W_1 of $\mathcal{W}$, then there is one axiom in the axiom set $\mathcal{A}$ initially. However, in extended incidence calculus, it is required that any formula which is a tautology should also be included in the axiom set. For this specific example, there are 7 formulae in $\mathcal{L}(P)$ which are regarded as tautologies as listed below, so they should all be included into set $\mathcal{A}$ and their incidence sets are $\mathcal{W}$.

r_1: Rainy$\vee$ Windy $\rightarrow$ Go to work
r_2: Go to work $\rightarrow$ Wear a suit
r_3: Sunny $\rightarrow$ Have a picnic
r_4: (Rainy$\vee$ Windy $\rightarrow$ Go to work) $\wedge$ (Go to work $\rightarrow$ Wear a suit)
r_5: (Rainy$\vee$ Windy $\rightarrow$ Go to work) $\wedge$ (Sunny $\rightarrow$ Have a picnic)
r_6: (Go to work $\rightarrow$ Wear a suit) $\wedge$ (Sunny $\rightarrow$ Have a picnic)
r_7: (Rainy$\vee$ Windy $\rightarrow$ Go to work) $\wedge$ (Go to work $\rightarrow$ Wear a suit) $\wedge$ (Sunny $\rightarrow$ Have a picnic)

The size of $\mathcal{A}$ is then increased from 1 to at least 15 after we extend $\mathcal{A}$ to be closed under $\wedge$. We denote this extended set as $\mathcal{A}^{\wedge}_{\top}$ ($\top$ stands for *true*) which is the set generated from $\mathcal{A}$ by adding tautologies into the set and then extend this set to be closed under $\wedge$.

Therefore, the final set of axioms is $\mathcal{A}^{\wedge}_{T} = \{$Rainy, $r_1, ..., r_7$, Rainy $\wedge r_1$, ..., Rainy $\wedge r_7$, $\}$. To obtain the bounds of incidence sets for a formula ϕ, it should involve $\mid \mathcal{A}^{\wedge}_{T} \mid$ steps implication checking between ϕ and an axiom ψ. For instance, if we let $\phi =$'Go to work', the lower bound of ϕ is then calculated based on equation (1) and $i_*(\phi) = \cup_{\psi \models \phi} i(\psi) = W_1$. Although it is required to carry out $\mid \mathcal{A}^{\wedge}_{T} \mid$ steps checking, in fact only the checking between axiom 'Rainy $\wedge r_1$' and ϕ is necessary and essential. All other steps are worthless to try.

However, if P is split into P_1, P_2, P_3 as stated in Section 1.2, a piece of evidence is propagated from P_1 to P_2 via the statements, then $\mid \mathcal{A}^{\wedge}_{T} \mid -1$ steps of worthless checking would be avoided. In another word, if $\mid \mathcal{L}(P) \mid \gg \Sigma_j \mid \mathcal{L}(P_j) \mid$, then splitting P into small sets would certainly increase the efficiency.

Set	$\mid \mathcal{A} \mid$	$\mid \mathcal{A}^{\wedge}_{\top} \mid$	Goal	Time
P	1	15	Go to Work	2460
P_1, P_2	1	1	Go to Work	90

Table 1: The test result for Example 1.1

Table 1 shows that when there is only one set P, it takes 2460 (secs) to infer the lower bound of *Go to work* while it only takes 90 (secs) to get the result if P is split.

Therefore for the first question above, the answer would be as follows.

Principle 1:*Assume that Q_1 and Q_2 are two basic questions*[1] *for which P_1 and P_2 contain the answers (propositions) for them respectively. If there are some logical implications between elements in P_1 and P_2, then keeping P_1 and P_2 as two sets of propositions would increase the efficiency and ignore meaningless inferences.*

This analysis also partially answers Question 2 above. That is, splitting a big set of propositions into smaller sets makes each of the relevant questions more obvious. To see whether it is always possible to make such division and whether such splitting is meaningful, we again reexamine the example in Section 2.

In Example 2.1, there are two propositions in set P, that is *grey*, *double* which can be put into two separate sets of propositions providing the answers for questions 'What colour is the coat?' and 'Is the coat double breasted?' respectively. However, there is no obvious implications between the elements in the two sets. For instance, the colour of a coat doesn't tell whether the coat is single-breasted or double-breasted. That is, when we generate the language set $\mathcal{L}(P)$ based on $P = \{grey, double\}$, we don't get any extra tautologies. So given a set of axioms $\mathcal{A}$, it is not necessary to extend the set before making any inferences. In this case, keeping P as a whole is more meaningful than splitting it into two but not quite coherent sets. Therefore the answer to the second question is stated in Principle 2.

Principle 2:*Assume that Q_1 and Q_2 are two basic questions for which P_1 and P_2 contain the answers (propositions) for them respectively. If there are no explicit logical implications between elements in P_1 and P_2, then constructing a set $P = P_1 \cup P_2$ as the set of propositions will not affect the efficiency of inference in general.*

6 Conclusion

In [4], it is proved that extended incidence calculus is equivalent to the Dempster-Shafer (DS) theory of evidence in representing evidence. That is, any information which can be represented in one theory can also be represented in the other. A frame of discernment in DS theory can be taken as a set of propositions in extended incidence calculus. Therefore, the system structure designed in [8] for DS theory could be partially adopted for the approach we proposed in this paper. However, a set of propositions in extended incidence calculus is not necessarily to be a frame of discernment all the time. Therefore, a fully implementation of the approach here is an extension of the system in [8].

Incidence calculus provides a mechanism for uncertainty reasoning by encoding probabilities on formulae indirectly. To obtain the probabilities or even the bounds of probabilities on formulae, it is necessary to calculate the incidence sets of these formulae first. This step could be very inefficient when a set of possible worlds is large or when a set of axioms is large. The main concern in this paper is to reduce the computational complexity by splitting a large language set into a number of small, but coherent sets and then propagate incidences from one set to another.

Future work is planned to further investigate the nature of splitting and to implement an automatic reasoning system based on the research.

[1] By *basic questions*, we mean that a question cannot be further divided into two or more sub-questions.

References

[1] A. Bundy, Incidence calculus: a mechanism for probabilistic reasoning, *J. Automated Reasoning*, Vol. 1, No.1, pp. 263-283, 1985

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

[3] Fagin,R. and J. Halpern, Uncertainty, belief and probability, IJCAI-89, pp. 1161-1167, 1989

[4] Liu,W. and A.Bundy, A comprehensive comparison between generalized incidence calculus and DS theory, *The Int. J. of Human-Computer Studies*, Vol. 40, pp, 1009-1032, 1994

[5] Liu,W., A.Bundy and D.Robertson, On the relations between incidence calculus and ATMS, *2nd European conference on symbolic and quantitative approaches to reasoning and uncertainty.* Lecture Notes in Computer Science 747:248-256 Springer.

[6] Liu,W. and A.Bundy, Constructing probabilistic ATMS using incidence calculus., To appear *International Journal of Approximate Reasoning*, 1996.

[7] Liu,W., *Extended incidence calculus and its comparison with related theories.* PhD thesis. Dept. of Artificial Intelligence, University of Edinburgh. 1995.

[8] Lowrance, J.D., Garvey,T.D. and Strat,T.M., A framework for evidential reasoning systems. *AAAI-86*, 896-903, 1986