國立交通大學

資訊管理研究所

博 士 論 文

採用模糊資訊的推理：方法與應用

Reasoning with Fuzzy Information:

Methods and Applications

研究生：高韓英

指導教授：黎漢林 博士

中華民國九十三年六月

採用模糊資訊的推理：方法與應用

Reasoning with Fuzzy Information: Methods and Applications

研 究 生：高　韓　英　　Student：Han-Ying Kao

指導教授：黎　漢　林　　Advisor：Han-Lin Li

國 立 交 通 大 學

資 訊 管 理 研 究 所

博 士 論 文

A Dissertation

Submitted to Institute of Information Management

College of Management

National Chiao Tung University

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Information Management

June 2004

Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

Dedication

To Dad, Mom, Leo and Athena,

who complete me.

採用模糊資訊的推理：方法與應用

學生：高韓英　　　　　　　　　　　　　　　　　　指導教授：黎漢林

國立交通大學資訊管理研究所博士班

摘　　　要

對於專家系統或是決策支援系統而言，推理是一項重要的工作。現實世界中，有三類常見的推理工作：預測、診斷、與規劃。在形形色色的知識庫與運算機制中，貝氏網路與影響圖是很普遍的圖形化模式，常用來處理不確定情況下的推理與決策。過去有許多學者提出各種演算法，試圖解決貝氏網路或影響圖上的查詢。然而，這些方法通常存在一些限制。首先，相關的參數或機率值，必須是確定而非模糊的。當決策或推理環境中，無法取得確定的知識，而只能取得不完整或是模糊的資訊時，推理工作將難以進行。其次，傳統貝氏網路的推理方法，難以考慮額外的限制。再者，不同的推理工作無法同時完成，例如診斷與決策之規劃。鑑於上述之限制，本論文擴展傳統的貝氏網路，而發展出一般性的貝氏網路，在這一般性的貝氏網路中，有幾個重要的組成集合：離散隨機節點之集合、連續隨機節點之集合、決策節點之集合，確定性參數之集合、與模糊參數之集合。除了傳統上只考慮離散隨機節點與確定參數的推理演算法，本論文研究三類貝氏網路的特殊題型，並提出解答的方法。這三類特殊推理題型為：(1) 考慮離散隨機節點與模糊參數之診斷，(2) 考慮離散隨機節點與模糊參數之診斷及決策，與(3) 考慮連續隨機節點之診斷與決策。

本論文的特色包含下列幾點：(1) 擴展傳統的貝氏網路，而發展出一般性的貝氏網路，其中考慮：離散隨機節點之集合、連續隨機節點之集合、決策節點之集合，確定性參數之集合、與模糊參數之集合。此一般性的貝氏網路，將作為本研究的基礎架構。(2) 解決在一般性貝氏網路上的模糊推理問題,包括牽涉模糊參數與可能性分配的題型。(3) 在推理的過程中，考慮無法納入正規知識庫的額外的限制或知識。(4) 在靜態與動態的環境下，解答針對貝氏網路的查詢。(5) 將發展的推理模式與方法，應用於醫療資訊或供應鏈管理的個案。所有的應用個案皆有詳細的解說。

關鍵詞: 模糊推理，貝氏網路，影響圖，供應鏈管理，醫療資訊。

Reasoning with Fuzzy Information: Methods and Applications

Student：Han-Ying Kao　　　　　　　　　　　　　　　　　　　Advisor：Han-Lin Li

Institute of Information Management
National Chiao Tung University

ABSTRACT

Reasoning is a major task to an expert system or a decision support system. Three types of reasoning tasks prevail in real-world applications: prediction, diagnosis and planning. Among the various knowledge bases and computation schema, Bayesian networks and influence diagrams are well-known graphical models for reasoning and decision-making under uncertainty. Many algorithms have been designed to answer the queries on a Bayesian network or an influence diagram. However, several limitations persist in the conventional methods. First, all relevant parameters are assumed to be crisp. Second, extra constraints or knowledge regarding belief propagation in Bayesian networks are difficult to embed. Third, diagnosis and planning cannot be completed in the same place. Motivated by the limitations mentioned above, this dissertation extend the traditional Bayesian networks to general Bayesian networks (GBN) that are composed of several components: the set of discrete random nodes, continuous random nodes, decision nodes, crisp parameters, and fuzzy parameters. In addition to the conventional reasoning problems that consider only crisp nodes and crisp parameters, three categories of reasoning are solved as the special cases (subsets) of general Bayesian networks: (1) diagnosis with discrete random nodes and fuzzy parameters; (2) diagnosis and decision-making with discrete random nodes and fuzzy parameters; and (3) diagnosis and decision-making with continuous random nodes in dynamic environments.

The distinguished features of this dissertation include: (1) extend the traditional Bayesian networks to general Bayesian networks, including discrete random nodes, continuous random nodes, decision nodes, crisp parameters, and fuzzy parameters. The general Bayesian networks are induced as the general research framework; (2) solve fuzzy reasoning tasks in three subsets of GBN where fuzzy parameters and possibility distributions are considered; (3) consider extra knowledge or constraints for the belief propagation, which are not implemented in the formal knowledge bases; (4) answer the queries from Bayesian networks in dynamic as well as static environments; (5) the reasoning models and methods are applied to the cases from medical informatics and supply chain management. All the applications are developed and illustrated in details.

Keywords: fuzzy reasoning, Bayesian networks, influence diagrams, supply chain management, medical informatics.

誌　　謝

能夠完成這本論文，將博士班學業告一段落，心中充滿無限的感恩。

首先，我想謝謝指導教授黎漢林老師，六年來對我的耐心教誨與指引，包容我這個平庸的學生，讓我有機會練習獨立研究，並且學著認真呵護、珍愛自己的每一篇作品。黎老師對研究的執著、與帶領學生的苦心，亦是我想努力學習的。若我在博士班期間，有任何學術上的成果，都要歸功於黎老師。

計畫書及論文口試期間，對於口試委員王小璠教授、游伯龍老師、曾國雄老師、陳茂生教授、溫于平教授的指正，感到完全地折服與受用。這段時間的收穫，比我預期的超出許多，也見識到老師們嚴謹治學、提攜後進的熱誠。您們將是我在日後學術生涯中，效法與趨近的目標。

一起用功的同學～菁蓉、念祖、榮發、嘉珍等，還有好多同研究室的學弟妹們，謝謝你們，讓我在交大得有歸屬感、學習的路上從不寂寞。

家人的支持與鼓勵，是我堅實的後盾。先生嘉輝總是陪在我身邊，分享我的笑聲與淚水，給我安全停泊的港口；博三時女兒悅珊的出生，為我們的生活帶來無比的歡愉，也讓我對女性的堅韌潛力，產生充分信心。最感謝的，是一直無條件欣賞我的爸爸高智雄先生、與媽媽劉純美女士～不管人生境遇如何，總是那麼愛我、相信我；還有我最可愛的親手足～姊姊岱伶、妹妹巧巧與岱琪、弟弟韓中，讓我在人生的歷程中，常有知音的陪伴。

最後，我想將博士學位的榮耀，獻給天國的兩位女性～我的祖母陳月琴女士(1914~1992) 與駱月裡女士(1915~1998)。她們的智慧、美麗、與慈愛，遠非我所能及，但卻選擇在上個世紀，將她們的一生奉獻給家庭。是她們讓我瞭解到女性的胸襟與偉大，教我懂得珍惜現代女性所掌握的幸福，也促使我更堅定這一生所追求的目標。

口試結束，當老師們向我恭賀時，我並沒有預期的狂喜，有的只是學而無涯的自省，與任重而道遠的自我鞭策。我會繼續努力的。

Contents

摘要 i

Abstract ii

誌謝 iii

Contents iv

Tables vi

Figures vii

Chapter 1 Introduction 1

1.1 Research background 1

1.2 Research Objectives and Framework 4

Chapter 2 Literatures review 10

2.1 Expert systems and probabilistic reasoning 10

2.2 Bayesian networks 12

2.3 Fuzzy sets and theory 14

Chapter 3 Diagnosis with fuzzy parameters 16

3.1 Reasoning with crisp information 16

3.2 Problem and goals 21

3.3 Model development 24

3.3.1 Fuzzy parameters 24

3.3.2 Fuzzy Abductive Models 26

3.4 Solution and illustrative examples 28

3.5 Discussions and conclusions 37

Chapter 4 Diagnosis and decision with fuzzy parameters 39

4.1 Influence diagrams 40

4.2 Problem and goals 47

4.3 Model development 49
4.4 Algorithm and solutions 53
4.5 Discussions and conclusions 56
Chapter 5 Diagnosis and decision with fuzzy nodes 59
5.1 Reasoning in supply chain management 59
5.2 Problem and Goals 61
5.2.1 Dynamic Bayesian networks with fuzzy nodes 64
5.2.2 Uncertainties in supply chains 64
5.3 Model development 65
5.4 Algorithms and solutions 74
5.4.1 Phase I: diagnostic phase 74
5.4.2 Phase II: Optimization phase 78
5.4.3 Linearization strategies 87
5.5 Discussions and conclusions 90
Chapter 6 Discussions and conclusions 93
6.1 Discussions 93
6.2 Future extensions 95
6.3 Concluding remarks 96
References 98

Tables

Table 1 The Associated Conditional Probability Distribution of Figure 1(b) 7

Table 2 The membership functions of fuzzy probabilities 22

Table 3 Solution of Example 2.1. 32

Table 4 The conditional probability distribution of Example 2.2. 34

Table 5 The membership functions of fuzzy probabilities in Example 2.2..................35

Table 6 Solution of Example 2.2. 38

Table 7 The conditional probabilities of pathogens, tests, and signs of UTI 44

Table 8 The conditional probabilities of Signs ($Sign_i$)...45

Table 9 Conditional probabilities of Coverage with resistance (*Resist* =1) 45

Table 10 The membership functions of fuzzy probabilities 50

Table 11 Solution of Example 3 58

Table 12 Description of nodes in the dynamic influence diagram in Figure 11.............69

Table 13 The probability/possibility distributions for the dynamic Influence Diagrams in Example 4. 71

Table 14a The results of simulation (λ -level = 0.0)...79

Table 14b The results of simulation (λ -level = 0.5)...80

Table 14c The results of simulation (λ -level = 1.0): 81

Table 15 Description of IT solutions 83

Table 16 Estimated costs/utilities of IT solutions in Example 4 (scale: 0 to 100) 84

Table 17 The possibility functions of the fuzzy parameters in Example 4. 85

Table 18 Solution report of Example 4...91

Figures

Figure 1 (a) an example of Bayesian networks, (b) the tree structure as clustering B and C into Z 6

Figure 2 Research framework of the dissertation 9

Figure 3 The membership function $\mu_{\tilde{x}_1}(x_1)$ of $\tilde{x}_1$ 20

Figure 4 The membership function $\mu_{\tilde{x}_7}(x_7)$ of $\tilde{x}_7$ 20

Figure 5 A membership function of fuzzy probability 24

Figure 6 A Bayesian network of the relationships between JITP techniques and performance measures 34

Figure 7 The input-output diagram of the optimization model for Chapter 4 42

Figure 8 A revised Bayesian network for Urinary tract infection 43

Figure 9 Research framework of Chapter 5 61

Figure 10 Cause-effect diagram of the two-echelon automotive supply chain 63

Figure 11 A dynamic influence diagram of the supply chain 67

Figure 11a Time expansion of the dynamic influence diagram in Figure 11 68

Figure 12 A schematic view of DSS 96

Chapter 1 Introduction

Reasoning algorithms is a core issue to an expert system or a decision support system. In many domains, such as medical inference or industrial informatics, there are at least three types of reasoning tasks in a decision support system: *prediction*, *diagnosis*, and *decision-making* [32,41]. To conduct the reasoning tasks, an expert system or decision support system needs a knowledge representation mechanism for the knowledge base. Bayesian networks are commonly used graphical probabilistic models for the knowledge base. This chapter will review the basics of expert systems and reasoning.

1.1 Research background

Expert systems are a kind of information systems which should be able to process an memorize information, learn and reason in both deterministic an uncertain situations, communicate with human and/or other expert systems, make appropriate decisions, and explains why these decision work.

Castillo et al [2] classified the problems that an expert system can deal with into two types: *deterministic* and *stochastic*. Deterministic problems can be formulated using a set of rules that relates several well-defined objects. Experts systems that deal with deterministic problems are known as *rule-based expert systems*. In stochastic or uncertain situations it is necessary to introduce some means for handling uncertainty, such as certainty factors, fuzzy logic, probability, and so on. Expert systems that use probability as a measure of uncertainty are know as probabilistic expert systems, and the strategy they use is know as *probabilistic reasoning* or *probabilistic inference*. The ability to use both predictive and diagnostic information is an important component of plausible reasoning, and improper handling of such information leads to strange results. So, Pearl [35] classified the patterns of plausible reasoning into abductive reasoning and inductive reasoning. Deduction, or prediction, is a logical process from a hypothesis to deduce evidence where probabilistic relationships are involved [35]. For example, if A is true, then B is true; that is A implies B. Abductive reasoning, or diagnosis, is a logical process that hypothetically explains experimental observations. For example, if A implies B, then finding B is true makes A more credible.

This dissertation will focus on the abductive reasoning and decision-making models in expert systems. In this dissertation, Bayesian networks and influence diagrams play a central role in the uncertainty formalism.

Bayesian networks [34,35] are directed acyclic graphs (DAG) in which the nodes represent the variables, the arcs represent the direct causal influences between the linked variables, and the strengths of these influences are expressed by forward conditional probabilities. The semantics of Bayesian networks demands a clear correspondence between the topology of a DAG and the dependency relationships portrayed by it. They are widely used knowledge representation and reasoning tools for various domains under uncertainty [1,2,4,8,13-18,20,23,27,34,35].

Influence diagrams are a special type of Bayesian networks with three kinds of nodes: decision nodes, chance nodes, and a value node. Decision nodes, shown as squares, represent choices available to the decision-makers. Chance nodes, shown as circles, represent random variables (or uncertain quantities). Finally, the value node, shown as a diamond, represents the objective (or utility) to be maximized. In a multiple objective decision making model, there may be more than one value nodes. There are two methods for determining the optimal decision policy from an influence diagram [35]. The first, proposed by Howard and Matheson [11], consists of converting the influence diagram to a decision tree and solving for the optimal policy within the tree, using exp-max labeling procedure. The second approach, proposed by Shachter, to decision-making in influence diagrams consists of eliminating modes from diagram through a series of value-preserving transformations.

Several methods have been developed for solving abductive or diagnostic reasoning problems in Bayesian networks. Exact methods exploit the independence structure contained in the network to efficiently propagate uncertainty [2,35]. Meanwhile, stochastic simulation methods provide an alternative approach suitable for highly connected networks, in which exact algorithms can be inefficient [35]. Recently, search-based approximate algorithms, which search for high probability configurations through a space of possible values, have emerged as a new alternative [36]. On the other hand, two key approaches have been proposed for symbolic inference in Bayesian networks, namely: the symbolic probabilistic inference algorithm (SPI) [38] and symbolic calculations based on slight modifications of standard numerical propagation algorithms [1,2].

The above methods have several limitations for reasoning from a Bayesian network or an influence diagram:

1. Most literatures focused on the discrete random nodes with discrete probability distributions.
2. All relevant parameters are assumed to be crisp.
3. Extra constraints or knowledge regarding belief propagation in Bayesian networks are difficult to embed.
4. Decision-making and diagnosis cannot be done in a complete model. Even in a compact graphical decision model, like influence diagrams, the proposed methods only focus on maximizing the expected gains but ignoring the problem diagnosis.

Those limitations restrict the usefulness of reasoning in Bayesian networks. First, the conditional probabilities between a random node and its parents could be fuzzy parameters because of the difficulties of learning accurately the causal relationships among the nodes. The decision makers may also feel awkward to make judgments for the linguistic vagueness or incomplete knowledge, which make the probability theory not suitable in problem formulation. Under such circumstances, the fuzzy nodes in a Bayesian networks can be introduced to overcome the obstacle. Additionally, knowledge workers often acquire additional information regarding inferences in Bayesian networks, particularly when facing diverse diagnostic scenarios. This information can relate to boundary, dependency or disjunctive conditions.

1.2 Research Objectives and Framework

Based on the limitations mentioned above, this dissertation is motivated to investigate and develop the reasoning methods for Bayesian networks and influence diagrams with improved features.

The objectives of this dissertation are as follow.

1. Develop the reasoning models that can contain various kinds of Bayesian networks that may include crisp discrete nodes, continuous nodes, crisp parameters, fuzzy parameters, and decision nodes.
2. Introduce extra knowledge or constraints into the reasoning models, which can perform the propagation more efficiently and effectively.
3. Design the model that can complete diagnosis and suggest optimal treatment simultaneously, which can facilitate the performance in a business or medical decision support systems.

For the common base of research, this dissertation first defines a general Bayesian networks as follow.

<u>Definition 1</u> General Bayesian networks.

A general Bayesian network (*GBN*) is a directed acyclic graph (DAG) representing the joint probability distribution of several sets of variables, including D*N*, *CN*, *XN*, *L*, *P*; that is .

GBN= (D*N*, *CN*, *XN, L, P*), where

DN denotes a set of discrete random nodes;

CN denotes a set of continuous random nodes;

P denotes a set of parameters (probabilities);

XN denotes the decision node set;

L denotes a set of directed links between the nodes, such that

$L=(DN,CN,XN) \times (DN,CN,XN)$ □

Based on the definition of GBN, we can induce several specific types of Bayesian networks. Consider a Bayesian network widely referred in Figure 1. Figure 1 represents the variables and their relationships from a medical problem. There are five random nodes, *A, B, C, D, E*. If all the random nodes in Figure 1 are discrete variables, and their probability distributions are crisp as in Table 1, then we can define a typical Bayesian network most common in the literatures, namely, $BN_1 = (DN, L, P)$.

If the parameters of the probability distributions are not crisp but fuzzy, for example, $P(+b|+a) = \tilde{x}_1$, $P(+b|-a) = \tilde{x}_2$, $P(+c|+a) = \tilde{x}_3$, $P(+c|-a) = \tilde{x}_4$, $P(+d|+b,+c) = \tilde{x}_5$, $P(+d|-b, +c) = \tilde{x}_6$, $P(+d|+b, -c) = \tilde{x}_7$, and $P(+d|-b, -c) = \tilde{x}_8$, then we can define the second type of Bayesian networks, BN_2 in the form of $BN_2 = (DN, L, \tilde{P})$, where the parameter set turns into fuzzy.

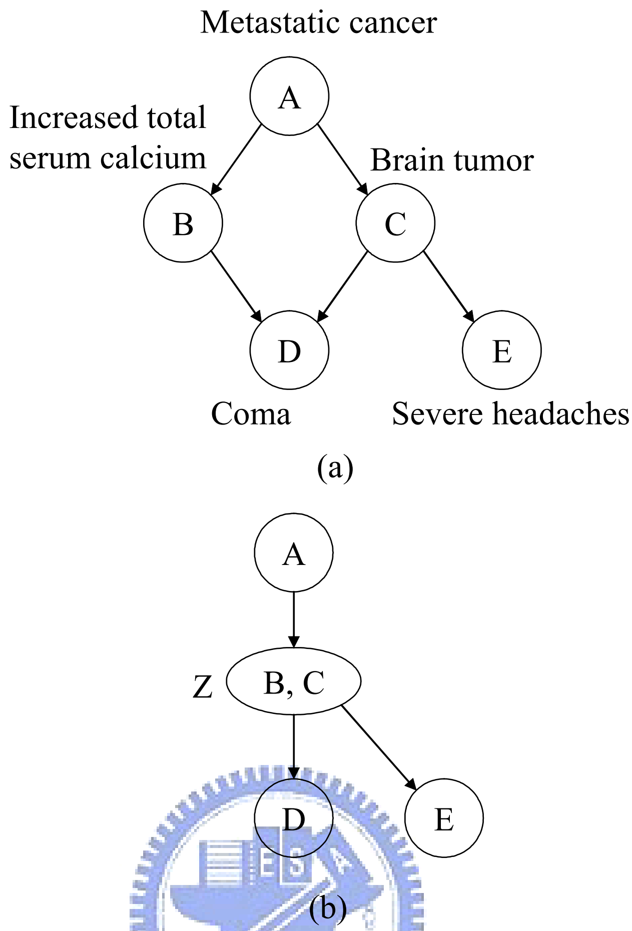

Figure 1: (a) an example of Bayesian networks, (b) the tree structure as clustering B and C into Z [35]

Furthermore, if the Bayesian networks involve not only discrete random nodes but also decision nodes, then the BN_2 can be extended into BN_3 in the form of BN_3= (*DN, XN, L,* $\widetilde{P}$), where the decision node set, *XN*, is added.

In many domains, there may be continuous variables involved. In such circumstances, the continuous random nodes must be added into the Bayesian networks, which induces the fourth type of Bayesian networks BN_4 in the form of BN_4= (*DN, CN, XN, L, P*), where the continuous random node set *CN* is included,

Table 1: The Associated Conditional Probability Distribution of Figure 1(b)

$P(+a) = 0.20$	
$P(+b\|+a) = 0.80$	$P(+b\|\text{-y}) = 0.20$
$P(+c\|+a) = 0.20$	$P(+c\|-a) = 0.05$
$P(+d\|+b, +c) = 0.80$	$P(+d\|-b, +c)= 0.80$
$P(+d\|+b, -c) = 0.80$	$P(+d\|-b, -c) = 0.05$
$P(+e\|+c) = 0.80$	$P(+e\|-c) = 0.60$

Additionally, a general Bayesian network is normally acyclic. However, in some special situations, the Bayesian networks may be cyclic. The feedback loops in cyclic Bayesian networks imply the time-series dependency between the network nodes, which consequently expend the static Bayesian networks into dynamic Bayesian networks [4].

After the Bayesian networks are constructed as the knowledge bases, the decision makers need to reason from the knowledge bases. This kind of reasoning tasks is called abductive reasoning. The general form of abductive reasoning is explained in the following.

Remark 1 Abductive reasoning.

Given a set of evidence or observations Ĕ from a *GBN*, define the set of unknown nodes Û⊂GBN\ Ĕ, the query of the belief (posterior) distribution of Û, *BEL*(Û| Ĕ), is an abductive reasoning problem. □

Since the conventional methods only answer very narrow scope of the queries on Bayesian networks, this dissertation develops several models to handle a set of specific reasoning problems in general Bayesian networks. In addition, these models are extended to consider the diagnosis and decision-making as well. Based on the four types of Bayesian networks introduced previously, there are four categories of reasoning problems discussed in this dissertation:

1. **Problem 1:** diagnosis with discrete random nodes and crisp parameters. This category is reasoning from the simplest type of the Bayesian networks, $BN_1 = (DN, L, P)$, and has been vastly studied in the literatures (Chapter 3).
2. **Problem 2:** diagnosis with discrete random nodes and fuzzy parameters in a static Bayesian network. This kind of problems is reasoning from $BN_2 = (DN, L, \widetilde{P})$ (Chapter 3).
3. **Problem 3:** diagnosis and decision-making with discrete random nodes and fuzzy parameters in a static influence diagram. This kind of problems is solved on $BN_3 = (DN, XN, L, \widetilde{P})$ (Chapter 4).
4. **Problem 4:** diagnosis and decision-making with continuous random nodes, decision nodes, and crisp parameters in a dynamic influence diagram. This type of problems is answered from $BN_4 = (DN, CN, XN, L, P)$ (Chapter 5).

For every category of problems, this dissertation first gives a description of problem formulation, and develops the reasoning model in a comprehensive and systematic way. Thereafter, the algorithms and solutions will be designed. One example or examples will be used to illustrate how to operate the reasoning methods, especially in medical informatics and supply chain systems. The outcomes and performances are examined carefully in the discussions. In the final chapter, some concluding remarks will be presented. The conceptual research framework and the dissertation structure are shown in Figure 2.

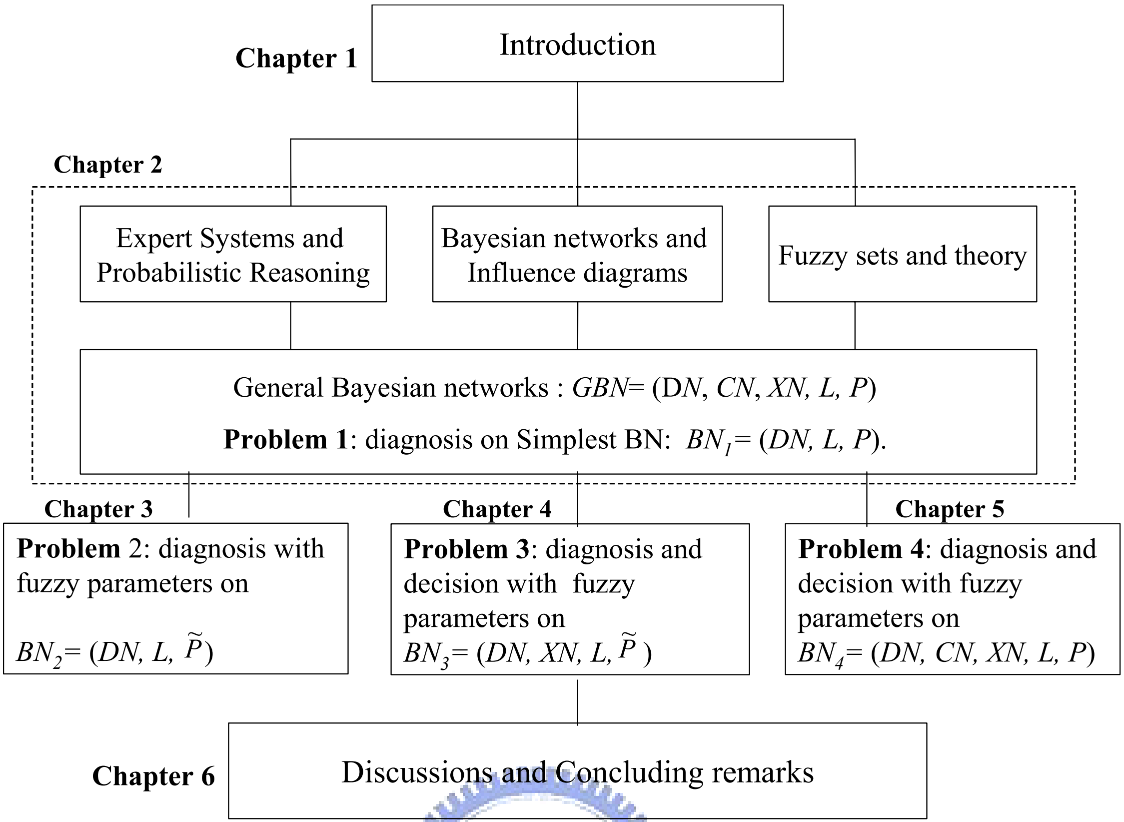

Figure 2: Research framework of the dissertation

Chapter 2 Literatures review

This chapter reviews the basic concepts of probabilistic reasoning, Bayesian networks, influence diagrams, and fuzzy sets.

2.1 Expert systems and probabilistic reasoning

First of all, this dissertation defines expert systems as follows.

Definition 2: **Expert systems**

An expert system can be defined as a computer system (hardware or software) that simulates human experts in a given area of specialization [2]. □

As such, an expert system should be able to process an memorize information, learn and reason in both deterministic an uncertain situations, communicate with human and/or other expert systems, make appropriate decisions, and explains why these decision work.

Castillo et al [2] classified the problems that an expert system can deal with into two types: *deterministic* and *stochastic*. Deterministic problems can be formulated using a set of rules that relates several well-defined objects. Experts systems that deal with deterministic problems are known as *rule-based expert systems*. In stochastic or uncertain situations it is necessary to introduce some means for handling uncertainty, such as certainty factors, fuzzy logic, probability, and so on. Expert systems that use probability as a measure of uncertainty are know as probabilistic expert systems, and the strategy they use is know as *probabilistic reasoning* or *probabilistic inference*. The ability to use both predictive and diagnostic information is an important component of plausible reasoning, and improper handling of such information leads to strange results. So, Pearl [35] classified the patterns of plausible reasoning into abductive reasoning and inductive reasoning. Deduction, or prediction, is a logical process from a hypothesis to deduce evidence where probabilistic relationships are involved [35]. For example, if A is true, then B is true; that is A implies B. Abductive reasoning, or diagnosis, is a logical process that hypothetically explains experimental observations. For example, if A implies B, then finding B is true makes A more credible.

George Polya [37] classified plausible reasoning into the following four:

1. **Inductive patterns**: "The verification of a consequence renders a conjecture more credible." For example, the conjecture "She didn't sleep well last night" becomes more credible when we verify, "She looks dispirited this morning".
2. **Successive verification of several consequences**: "The verification of a new consequence counts more or less if the new consequence differs more less from the former, verified consequences." For example, if in trying substantiating the conjecture "All ravens are black," we observe *n* Australian ravens, all of them black, our subsequent confidence in the conjecture will be increased substantially of the ($n+1$)-th ravens is a black Brazilian rather than another Australian ravens.
3. **Verification of improbable consequences**: "The verification of a consequence counts more or less according as the consequence is more or less improbable in itself." For example, the conjecture "She didn't sleep well" obtains more support from "She is nodding this morning" than from the more common observation "She looks dispirited this morning".
4. **Inference from analogy**: "A conjecture becomes more credible when an analogous conjecture turns out to be true." For example, the conjecture "Of all objects displacing the same volume, the sphere has the smallest surface" becomes more credible when we prove the relative theorem "Of all curves enclosing the same area, the circle has the shortest perimeter."

This dissertation will focus on the abductive reasoning and decision-making models in expert systems. In this research, Bayesian networks and influence diagrams play a central role in the uncertainty formalism.

2.2 Bayesian networks

Bayesian networks [34,35] are directed acyclic graphs (DAG) in which the nodes represent the variables, the arcs represent the direct causal influences between the linked variables, and the strengths of these influences are expressed by forward conditional probabilities. A simple example is given in Figure 1(b). The semantics of Bayesian networks demands a clear correspondence between the topology of a DAG and the dependency relationships portrayed by it. They are widely used knowledge representation and reasoning tools for various domains under uncertainty [1,2,4,8,13-18,20,23,27,34,35].

Influence diagrams are a special type of Bayesian networks with three kinds of nodes: decision nodes, chance nodes, and a value node. Decision nodes, shown as squares, represent choices available to the decision-makers. Chance nodes, shown as circles, represent random variables (or uncertain quantities). Finally, the value node, shown as a diamond, represents the objective (or utility) to be maximized. In a multiple objective decision making model, there may be more than one value nodes. There are two methods for determining the optimal decision policy from an influence diagram [35]. The first, proposed by Howard and Matheson, consists of converting the influence diagram to a decision tree and solving for the optimal policy within the tree, using exp-max labeling procedure. The second approach, proposed by Shachter, to decision-making in influence diagrams consists of eliminating modes from diagram through a series of value-preserving transformations.

Several methods have been developed for solving abductive or diagnostic reasoning problems in Bayesian networks. Exact methods exploit the independence structure contained in the network to efficiently propagate uncertainty [2,35]. Meanwhile, stochastic simulation methods provide an alternative approach suitable for highly connected networks, in which exact algorithms can be inefficient [35]. Recently, search-based approximate algorithms, which search for high probability configurations through a space of possible values, have emerged as a new alternative [36]. On the other hand, two key approaches have been proposed for symbolic inference in Bayesian networks, namely: the symbolic probabilistic inference algorithm (SPI) [38] and symbolic calculations based on slight modifications of standard numerical propagation algorithms [1,2].

The above methods have several limitations for reasoning from a Bayesian network or an influence diagram:

1. All network nodes or variables must be crisp.
2. All relevant parameters are assumed to be crisp.
3. Extra constraints or knowledge regarding belief propagation in Bayesian networks are difficult to embed.
4. Decision-making and diagnosis cannot be done in a complete model. Even in a compact graphical decision model, like influence diagrams, the proposed methods only focus on maximizing the expected gains but ignoring the problem diagnosis.

Those limitations restrict the usefulness of reasoning in Bayesian networks. First, the conditional probabilities between a node and its parents could be fuzzy parameters because of the difficulties of learning accurately the causal relationships among the nodes. The decision makers may also feel awkward to make judgments for the linguistic vagueness or incomplete knowledge, which make the probability theory not suitable in problem formulation. Under such circumstances, the fuzzy nodes in a Bayesian networks can be introduced to overcome the obstacle. Additionally, knowledge workers often acquire additional information regarding inferences in Bayesian networks, particularly when facing diverse diagnostic scenarios. This information can relate to boundary, dependency or disjunctive conditions.

2.3 Fuzzy sets and theory

Fuzzy sets were introduced by Zadeh [43] in 1965 to manipulate data and information processing uncertainties which statistics is not proper for use. It was particularly designed to mathematically represent uncertainty as well as vagueness and to offer formalized tools for handling the imprecision intrinsic to many domains.

Fuzzy sets are a means of representing and manipulating information not precise. A fuzzy subset $\tilde{A}$ of a set X can be can be defined as a set of ordered pairs, each with the first element from X and the second element from the interval [0,1], with exactly one ordered pair for each element of X. This defines a mapping as below.

$$\mu_{\tilde{A}} : X \to [0,1],$$

between elements of the set X and values in the interval [0,1]. The value zero is to represent complete non-membership, the value one is to represent complete membership, and values in between are to represent intermediate degrees of membership. The set X is referred as the universe of discourse for the fuzzy subset $\tilde{A}$. Usually, the mapping $\mu_{\tilde{A}}$ is described as a function, the membership function of $\tilde{A}$. The degree to which the statement "x is in $\tilde{A}$" is true is determined by finding the ordered pair (x, $\mu_{\tilde{A}}$). The degree of the statement to be true is the second element of the ordered pair.

<u>Definition 3</u>: Fuzzy membership functions.

Let X be a nonempty set. A fuzzy set $\tilde{A}$ in X is characterized by its membership function

$$\mu_{\tilde{A}} : X \to [0,1],$$

and $\mu_{\tilde{A}}$ is interpreted as the degree of membership of element x in fuzzy set $\tilde{A}$ for each x □

Visibly, $\tilde{A}$ is completely determined by the following expression.

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) \mid x \in X\}.$$

By the above expression, the terms membership function and fuzzy subset are used interchangeably. A fuzzy subset $\tilde{A}$ of a classical set X is called *normal* if there exists an $x \in X$ such that $\tilde{A}(x) = 1$. Otherwise, $\tilde{A}$ is subnormal. An α-level set (or α-cut) of a fuzzy set $\tilde{A}$ of X is a non-fuzzy set denoted by $[\tilde{A}]^{\alpha}$ and defined by

$$[\tilde{A}]^{\alpha} = \begin{cases} \{x \in X \mid \tilde{A}(x) \geq \alpha\}, & \text{if } \alpha > 0 \\ \text{cl}(\text{supp } \tilde{A}), & \text{if } \alpha = 0 \end{cases}$$

where $\text{cl}(\text{supp } \tilde{A})$ denotes the closure of the support of $\tilde{A}$. A fuzzy set $\tilde{A}$ of X is called convex if $[\tilde{A}]^{\alpha}$ is a convex subset of X for all $\alpha \in [0,1]$.

Similarly, a fuzzy number can be defined as follow.

Definition 4: Fuzzy numbers

A fuzzy number $\tilde{A}$ is a fuzzy set of the real line with a normal, fuzzy convex and continuous membership function satisfying the limit conditions, and $\lim_{t \to -\infty} \tilde{A}(t) = 0$. □

Based on the concepts reviewed in this chapter, next chapter will show how this dissertation solves the fuzzy reasoning problems on Bayesian network.

Chapter 3 Diagnosis with fuzzy parameters

Based on the review in Chapter 2, we understand that current adductive reasoning methods can solve very limited scope of the reasoning from Bayesian networks. This chapter will first illustrate the steps to solve a traditional abductive reasoning query, and then develop the model for diagnosis with crisp nodes and fuzzy parameters in Bayesian network.

3.1 Reasoning with crisp information

In this section, a simplest form of abductive reasoning is introduced as follow.

Problem 1: Given the evidence set Ě from $BN_1 = (DN, L, P)$, compute the belief distribution of $\hat{U} \subset BN_1 \setminus \check{E}$, $BEL(\hat{U}|\check{E})$. ■

Problem 1 is interpreted with the following case from medicine and **Example 1**. Consider the following example from Pearl [35].

"Metastatic cancer is a possible cause of a brain tumor and is an explanation for increased total serum calcium. Either of these could explain a patient falling into a coma. Severe headache is also possibly associated with a brain tumor."

Figure 1(b) shows a Bayesian network representing the above cause and effect relationships. Table 1 lists the causal influences in terms of conditional probability distributions. Each variable is characterized by the probability given the state of its parents. For instance: $C \in \{1,0\}$ represents the dichotomy between having a brain tumor and not having one, +c denotes the assertion C = 1 or "Brain tumor is present", and –c is the negation of +c, namely, C = 0. The root node, A, which has no parent, is characterized by its prior probability distribution. The above information can be used to solve the following reasoning problems.

<u>Example 1</u>: Compute the posterior probability of every A, B, and C, given the conditional probabilities in Table 1, and a situation involving a patient who is suffering from a severe headache (E=1) but has not fallen into a coma (D=0); that is, compute P(a|-d, +e), P(b|-d, +e) and P(c|-d, +e). □

Now this section reviews one conventional method, clustering, for computing the posterior probabilities with crisp parameters and no extra constraints. Consider the Bayesian network in Figure 1(b) with the crisp information in Table 1. Clustering [2,35] can transform Figure 1(b) into an equivalent tree structure in Figure 1(c), where nodes B and C are collapsed into a compound node $Z = B \& C$. Let $Z = \{z_1, z_2, z_3, z_4\}$ be a set of cardinalities of Z and $z_1 = (+b,+c)$, $z_2 = (-b,+c)$, $z_3 = (+b,-c)$, and $z_4 = (-b,-c)$. Moreover, let W_Y denote the state of all variables except Y; for example, W_A={$(z_1,-d+e)$, $(z_2,-d+e)$, $(z_3,-d+e)$, $(z_4,-d+e)$}. From Pearl [35], the value of $P(y|W_Y)$, which is the distribution of y conditioned on the value W_Y, can be calculated as below considering every instance of y.

$$
\left.\begin{aligned}
P(+a \mid W_A) &= \alpha_A P(+a) \sum_{i=1}^{4} P(z_i \mid +a) P(-d \mid z_i) P(+e \mid z_i) \\
P(-a \mid W_A) &= \alpha_A P(-a) \sum_{i=1}^{4} P(z_i \mid -a) P(-d \mid z_i) P(+e \mid z_i) \\
P(+b \mid W_B) &= \alpha_B \sum_{a=0}^{1} [P(a) \sum_{i=1,3} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] \\
P(-b \mid W_B) &= \alpha_B \sum_{a=0}^{1} [P(a) \sum_{i=2,4} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] \\
P(+c \mid W_C) &= \alpha_C \sum_{a=0}^{1} [P(a) \sum_{i=1,2} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] \\
P(-c \mid W_C) &= \alpha_C \sum_{a=0}^{1} [P(a) \sum_{i=3,4} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)]
\end{aligned}\right\} \qquad (1)
$$

where α_A, α_B, and α_C are the normalizing constant ensuring that

$$
\begin{aligned}
P(+a \mid W_A) + P(\text{-}a \mid W_A) &= 1 \\
P(+b \mid W_B) + P(\text{-}b \mid W_B) &= 1 \\
P(+c \mid W_C) + P(\text{-}c \mid W_C) &= 1
\end{aligned} \qquad (2)
$$

From (2), then intuitively

$$
\alpha = \alpha_A = \alpha_B = \alpha_C = \frac{1}{\sum_{a=0}^{1} P(a) \sum_{i=1}^{4} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)} \qquad (3)
$$

and

$$
\alpha \sum_{a} \sum_{z_i} P(a) P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i) = 1 \qquad (4)
$$

The value of $P(+a \mid W_A)$ in (1) is obtained below for the data in Table 1:

$$
\begin{aligned}
P(+a \mid W_A) = \alpha(.2)[&(.8)(.2)(1-.8)(.8) + (1-.8)(.2)(1-.8)(.8) \\
&+ (.8)(1-.2)(1-.8)(.6) + (1-.8)(1-.2)(1-.05)(.6)].
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
P(-a \mid W_A) = \alpha(1-.2)[&(.2)(.05)(1-.8)(.8) + (1-.2)(.05)(1-.8)(.8) \\
&+ (.2)(1-.05)(1-.8)(.6) + (1-.2)(1-.05)(1-.05)(.6)].
\end{aligned}
$$

From (1) and (3), then $\alpha = 2.432$, $P(+a \mid W_A)$=0.097, and $P(-a \mid W_A)$=0.903.

The answers to Example 1 are

$P(a \mid -d,+e) = (0.097, 0.903)$, $P(b \mid -d,+e) = (0.097, 0.903)$, $P(c \mid -d,+e) = (0.031, 0.969)$.

■

Observing the solution stated above, several limitations persist in the conventional reasoning methods.

First, all network nodes and relevant parameters are assumed to be crisp. This narrows the usefulness of reasoning methods when some parameters are hard to estimate. Freeling [7] claimed fuzzy probability as an extension of probability theory, which is more promising than possibility and probability theory as a decision aid. Second, extra constraints or knowledge regarding belief propagation in Bayesian networks are difficult to embed. Third, different reasoning tasks, such as diagnosis as well as treatment planning, cannot be completed in the same place. Those attributes are often needed in both business and medical informatics. Furthermore, the limitations encumber reasoning to be automated.

For some systematic or technical reasons, the conditional probabilities of the network nodes may be fuzzy, instead of crisp. For instance, $P(+b \mid +a)$ cannot be 0.8 but rather is a fuzzy number, say $\tilde{x}_1$, where $P(+b \mid +a) = \tilde{x}_1$, and is associated with a membership function $\mu_{\tilde{x}_1}(x_1)$, represented as follows. (See Figure 3)

$$\mu_{\tilde{x}_1}(x_1) = 5(x_1 - 0.6) - 5(|x_1 - 0.8| + x_1 - 0.8), \qquad 0.6 \le x_1 \le 1$$

where "$|*|$" denotes the absolute value of a term *.

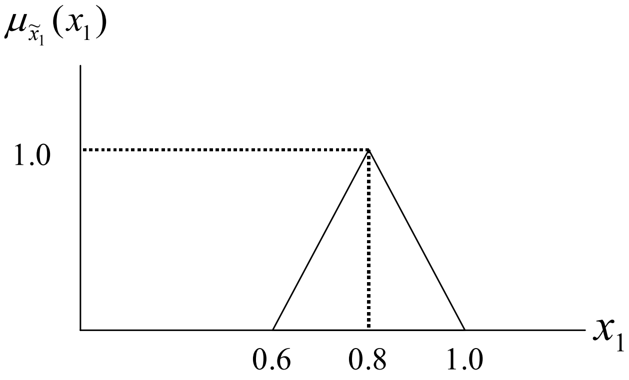

Figure 3: The membership function $\mu_{\tilde{x}_1}(x_1)$ of $\tilde{x}_1$

The above expression and Figure 3 mean that the domain of $\tilde{x}_1$ is between 0.6 and 1.0. If x_1=0.8 then $\mu_{\tilde{x}_1}(x_1)$=1, implying that x_1=0.8 is the most possible situation. If $x_1 \leq 0.6$ or $x_1 \geq 1$ then $\mu_{\tilde{x}_1}(x_1)$=0, the least possible manifestation of x_1. If x_1=0.7, then $\mu_{\tilde{x}_1}(x_1)$=0.5.

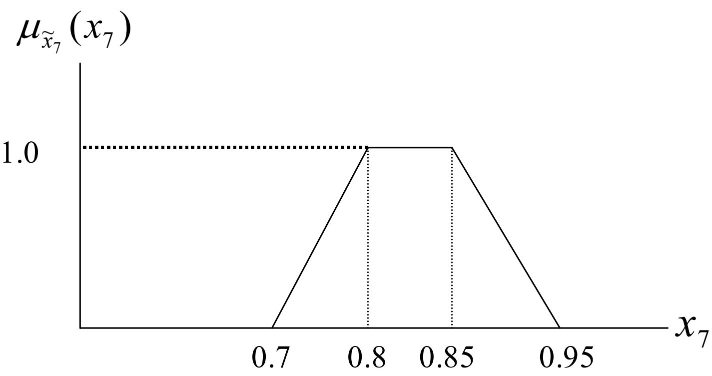

Figure 4: The membership function $\mu_{\tilde{x}_7}(x_7)$ of $\tilde{x}_7$

Fuzzy membership functions can be expressed in various ways. For example, let *P(+d|+b,+c) =* $\tilde{x}_7$ and express $\mu_{\tilde{x}_7}(x_7)$ as the following function (Figure 4).

$$\mu_{\tilde{x}_7}(x_7) = 10(x_7 - 0.7) - 5(|x_7 - 0.8| + x_7 - 0.8) - 5(|x_7 - 0.85| + x_7 - 0.85), \quad 0.7 \le x_7 \le 0.95 \quad .$$

$\mu_{\tilde{x}_7}(x_7)$ is a trapezoid membership function and comprises four line segments, where $0.8 \le x_7 \le 0.85$ has the maximal membership.

3.2 Problem and goals

This chapter discusses reasoning with crisp nodes and fuzzy parameters as the following problem.

<u>Problem 2</u>: Given the evidence set Ĕ from BN_2= ($DN, L, \tilde{P}$), compute the belief distribution of Û⊂BN_2\ Ĕ, *BEL*(Û| Ĕ).■

The fuzzy parameters are denoted by as $\tilde{x}_i$, $i = 1,2,\ldots,8$, where P(+b|+a) = $\tilde{x}_1$, P(+b|-a) = $\tilde{x}_2$, P(+c|+a) = $\tilde{x}_3$, P(+c|-a) = $\tilde{x}_4$, P(+d|+b,+c) = $\tilde{x}_5$, P(+d|-b, +c) = $\tilde{x}_6$, P(+d|+b, -c) = $\tilde{x}_7$, and P(+d|-b, -c) = $\tilde{x}_8$. Table 2 lists the membership functions of the fuzzy parameters, among which $\mu_{\tilde{x}_7}(x_7)$ and $\mu_{\tilde{x}_8}(x_8)$ are trapezoid membership functions while the remainder are triangular functions.

After introducing the fuzzy probabilities, the **Example 1** turns into a more complex problem as **Example 2**.

Table 2: The membership functions of fuzzy probabilities

Parameter $\tilde{x}_i$	$\mu_{\tilde{x}_i}(x_i)$	Domain of x_i
$P(+b \mid +a) = \tilde{x}_1$	$5(x_1 - 0.6) - 5(\lvert x_1 - 0.8\rvert + x_1 - 0.8)$	[0.6,1.0]
$P(+b \mid -a) = \tilde{x}_2$	$10(x_2 - 0.1) - 10(\lvert x_2 - 0.2\rvert + x_2 - 0.2)$	[0.1,0.3]
$P(+c \mid +a) = \tilde{x}_3$	$10(x_3 - 0.1) - 15(\lvert x_3 - 0.2\rvert + x_3 - 0.2)$	[0.1,0.25]
$P(+c \mid -a) = \tilde{x}_4$	$25(x_4 - 0.01) - 17.5(\lvert x_4 - 0.05\rvert + x_4 - 0.05)$	[0.01,0.15]
$P(+d \mid z_1) = \tilde{x}_5$	$5(x_5 - 0.6) - 5(\lvert x_5 - 0.8\rvert + x_5 - 0.8)$	[0.6,1.0]
$P(+d \mid z_2) = \tilde{x}_6$	$10(x_6 - 0.7) - 10(\lvert x_6 - 0.8\rvert + x_6 - 0.8)$	[0.7,0.9]
$P(+d \mid z_3) = \tilde{x}_7$	$10(x_7 - 0.7) - 5(\lvert x_7 - 0.8\rvert + x_7 - 0.8) - 5(\lvert x_7 - 0.85\rvert + x_7 - 0.85)$	[0.7,0.95]
$P(+d \mid z_4) = \tilde{x}_8$	$25(x_8 - 0.01) - 12.5(\lvert x_8 - 0.05\rvert + x_8 - 0.05) - 25(\lvert x_8 - 0.07\rvert + x_8 - 0.07)$	[0.01, 0.09]

Example 2: Compute the belief distributions *P(a|-d, +c)*, *P(b|-d, +c)*, and *P(c|-d, +c)*, given the fuzzy membership functions in Table 2 and some constraints related to belief propagation.

Current abductive reasoning methods have difficulties in solving **Problem 2** and Example 2 since it involves fuzzy information and extra constraints.

Consider abductive reasoning with constraints. For a given Bayesian network, knowledge workers (such as clinicians) may have professional judgments regarding the features of certain nodes and the relationships among them in particular diagnostic backgrounds. These features and relationships can take the form of various constraints [26].

1. Boundary constraints:

   From additional information or observations, clinicians can infer that the posterior probability of A given E=1 and D=0 should be higher than 0.1 but lower than 0.3, which is expressed as

$$0.1 \leq P(+a\,|\,\text{-}d,+e) \leq 0.3 \tag{5}$$

2. Functional dependency:

   The beliefs of certain nodes are functionally dependent. For example, clinicians can judge that the posterior probability of B is roughly a certain multiple of that of A given E=1 and D=0, which is expressed as

$$P(+a\,|\,\text{-}d,+e) \leq 2P(+b\,|\,\text{-}d,+e) \tag{6}$$

3. Disjunctive constraints:

   Sometimes disjunction may occur between nodes. For example, a doctor may estimate that either $P(+a\,|\,\text{-}d,+e)$ or $P(+b\,|\,\text{-}d,+e)$ is equal to or below 0.2, which is expressed as

$$\text{Either } P(+a\,|\,\text{-}d,+e) \leq 0.2 \text{ or } P(+b\,|\,\text{-}d,+e) \leq 0.2 \tag{7}$$

By introducing these constraints into the reasoning system, the following problems are formulated.

Example 2.1: Compute the belief distributions $P(a|\text{-}d, +e)$, $P(b|\text{-}d, +e)$, and $P(c|\text{-}d, +e)$, given the fuzzy membership functions in Table 2 and the following constraints.

$0.1 \leq P(+a\,|\,\text{-}d,+e) \leq 0.3$,

$P(+b\,|\,\text{-}d,+e) \leq 2P(+c\,|\,\text{-}d,+e)$

Either $P(+a\,|\,\text{-}d,+e) \leq 0.2$ or $P(+b\,|\,\text{-}d,+e) \leq 0.2$.

Example 2.1 is more complicated and difficult than Example 1 when solved using current propagation methods.

3.3 Model development

The following illustrates another approach for calculating the posterior probabilities with fuzzy parameters.

3.3.1 Fuzzy parameters

Consider a membership function $\mu_{\tilde{x}}(x)$ of $\tilde{x}$, as displayed in Figure 5. This piecewise linear function generally is expressed as

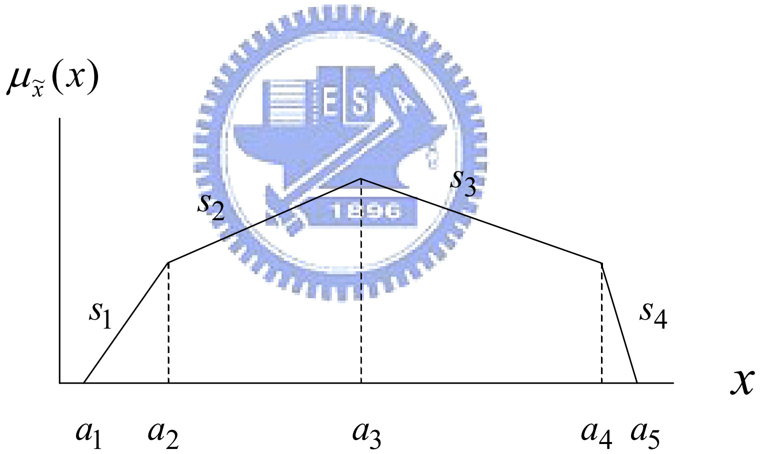

Figure 5: A membership function of fuzzy probability

$$\mu_{\tilde{x}}(x)=\begin{cases} s_1(x-a_1), & a_1 < x \le a_2 \\ \mu(a_2)+s_2(x-a_2), & a_2 < x \le a_3 \\ \mu(a_3)+s_3(x-a_3), & a_3 < x \le a_4 \\ \mu(a_4)+s_4(x-a_4), & a_4 < x \le a_5 \\ 0, & elsewhere. \end{cases} \quad (8)$$

Computing the above expression is complex. Consequently, this work employs an efficient method of expressing a piecewise linear function. Consider the following proposition.

<u>Proposition 1</u> Let $\mu_{\tilde{x}}(x)$ denote the membership function of fuzzy variable $\tilde{x}$, as displayed in Figure 4, where $a_j, j=1,2,...,m$ represent the break points of $\mu_{\tilde{x}}(x)$, and $s_j, j=1,2,...,n$ are the slopes of line segments between a_j and a_{j+1}, and $\mu_{\tilde{x}}(x)$ is the sum of absolute terms [24,40]:

$$\mu_{\tilde{x}}(x) = \mu(a_1) + s_1(x - a_1) + \sum_{j=2}^{m} \frac{s_i - s_{i-1}}{2}(|x - a_j| + x - a_j) \tag{9}$$

If $\mu_{\tilde{x}}(x)$ in (9) is to be maximized, then the following proposition is used for convenient linearization.

<u>Proposition 2</u> Maximizing a function $\mu_{\tilde{x}}(x)$ in (9) requires solving the following linear program [24,40]:

$$\left.\begin{array}{l}
Max \qquad z = s_1(x-a_1) + 2\sum_{j=2}^{m} \frac{s_i - s_{i-1}}{2}(x - a_j + \sum_{k=1}^{j} d_k) \\
subject \ \ to \\
\qquad x + d_1 \ge a_2, \\
\qquad x + d_1 + d_2 \ge a_3, \\
\qquad \qquad \vdots \\
\qquad x + d_1 + d_2 + \cdots + d_{m-1} \ge a_m, \\
\qquad 0 \le d_1 \le a_2, \\
\qquad 0 \le d_{k-1} \le a_k - a_{k-1}, \quad for \quad k = 2,3,\cdots,m, \\
\qquad x \in F(feasible \quad set).
\end{array}\right\} \tag{10}$$

Proof:

Since $d_{k-1} \le a_k - a_{k-1}$, then clearly

$x \ge a_k - (d_1 + d_2 + \cdots + d_{k-1}) \ge a_{k-1} - (d_1 + d_2 + \cdots + d_{k-2})$, so constraint

$x + d_1 + d_2 + \cdots + d_{k-2} \geq a_{k-1}$ is converted by constraint

$x + d_1 + d_2 + \cdots + d_{k-2} + d_{k-1} \geq a_k$, *for* $k = 2,3,\cdots,m$. ■

From Proposition 2, the non-linear membership functions are transformed into equivalent linear functions.

3.3.2 Fuzzy Abductive Models

To compute the belief distribution of the unknown nodes in a Bayesian networks with fuzzy parameters, there are several alternative objective functions. Consider **Example 2.1**.

1. Estimate the upper/ lower bound for P(+a|-d, +e), P(+b|-d, +e), P(+c|-d, +e) by maximizing/ minimizing the beliefs, respectively. e.g.

   Maximize P(+a|-d, +e) → Upper bound of P(+a|-d, +e),

   Minimize P(+a|-d, +e) → Lower bound of P(+a|-d, +e).

2. Generate a pair of belief, e.g. (P(+a|-d, +e)min μ, P(+a|-d, +e)max μ) with respect to the maximal/ minimal confidence for fuzzy parameters. e.g.

   Maximize $\mu_{\tilde{x}_i}(x_i)$ → under maximal confidence for fuzzy parameters.

   Minimize $\mu_{\tilde{x}_i}(x_i)$ → under minimal confidence for fuzzy parameters.

3. Generate the distributions of P(+a|-d, +e), P(+b|-d, +e), P(+c|-d, +e) by α-cut and fuzzy simulation.

All the above classes of the objectives can be implemented based on the decision-makers' needs or preferences. This dissertation chooses the second class as the objectives.

Since there are several fuzzy parameters involved in Problem 2, this dissertation will estimate the belief distribution for the unknown nodes with the maximal and minimal confidence. The belief distribution under maximal confidence will be estimated by maximizing the fuzzy membership functions; oppositely, the belief distribution under minimal confidence will be estimated by minimizing the fuzzy membership function.

Building upon the clustering method, Proposition 1 and 2, the abductive model for solving Example 2.1 is formulated below.

Model 1(a) (for maximal confidence)

$$\begin{aligned} &\textit{Maximize} \quad \mu_{\tilde{x}_i}(x_i), \quad i = 1,2,\cdots,8, \\ &\textit{subject to} \quad (1), \\ &\qquad 0.1 \le P(+a \mid \text{-}d,+e) \le 0.3, \\ &\qquad P(+b \mid \text{-}d,+e) \le 2P(+c \mid \text{-}d,+e), \\ &\qquad \textit{Either} \quad P(+a \mid \text{-}d,+e) \le 0.2 \quad \textit{or} \quad P(+b \mid \text{-}d,+e) \le 0.2, \end{aligned} \tag{11}$$

Model 1(b) (for minimal confidence)

$$\begin{aligned} &\textit{Minimize} \quad \mu_{\tilde{x}_i}(x_i), \quad i = 1,2,\cdots,8, \\ &\textit{subject to} \quad (1), \\ &\qquad 0.1 \le P(+a \mid \text{-}d,+e) \le 0.3, \\ &\qquad P(+b \mid \text{-}d,+e) \le 2P(+c \mid \text{-}d,+e), \\ &\qquad \textit{Either} \quad P(+a \mid \text{-}d,+e) \le 0.2 \quad \textit{or} \quad P(+b \mid \text{-}d,+e) \le 0.2, \end{aligned} \tag{12}$$

where the objective function maximize and minimize all fuzzy membership functions. Since

(4) contains numerous non-separate nonlinear terms, **Model 1** is a highly non-linear and nonconvex program. This dissertation will deal with the disjunctive constraint first and takes care of the nonlinear issue in the following proposition.

Proposition 3 A disjunctive constraint $f(\tilde{x}) \le 0$ or $g(\tilde{x}) \le 0$ can be expressed by the following inequalities.

$$\left.\begin{aligned} & M(\theta_1 - 1) \le f(\tilde{x}) \le M\theta_1 + M(1-\theta_2), \\ & M(\theta_2 - 1) \le g(\tilde{x}) \le M\theta_2 + M(1-\theta_1) \\ & \varepsilon \le \theta_2 + \theta_1 \le 1. \end{aligned}\right\} \tag{13}$$

where θ_1 and θ_2 are 0-1 variables, M is a relatively large number, and ε is a relatively small positive number.

The four possible combinations of θ_1 and θ_2 can be checked as follows: (i) for θ_1= 1, θ_2=1 the constraints are $0 \le f(\tilde{x}) \le M$ and $0 \le g(\tilde{x}) \le M$, which are inactive constraints; (ii) for θ_1= 0, θ_2=1 then $-M \le f(\tilde{x}) \le 0$ and $0 \le g(\tilde{x}) \le 2M$, meaning that when $g(\tilde{x}) \ge 0$, $f(\tilde{x})$ must be 0 or less; (iii) for θ_1= 1, θ_2=0, the constraints are $0 \le f(\tilde{x}) \le 2M$ and $-M \le g(\tilde{x}) \le 0$, which implies that when $f(\tilde{x}) \ge 0$, $g(\tilde{x})$ must be 0 or less; (iv) for θ_1= 0, θ_2=0 the constraints become $-M \le f(\tilde{x}) \le M$ and $-M \le g(\tilde{x}) \le M$, which are inactive constraints. The third constraint in (13) excludes the combinations θ_1= 1, θ_2=1 and θ_1= 0, θ_2=0. To summarize, (13) implies that either $f(\tilde{x}) \le 0$ or $g(\tilde{x}) \le 0$ must be satisfied.

3.4 Solution and illustrative examples

Abductive reasoning problems in certain applications are solved below using the proposed constrained optimization approach.

Example 2.1 is solved using the following program.

$$Maximize\quad \mu_{\tilde{x}_i}(x_i),\ \ i=1,2,\cdots,8,\ \ \text{(for the maximal confidence), or}$$

$$Minimize\quad \mu_{\tilde{x}_i}(x_i),\ \ i=1,2,\cdots,8,\ \ \text{(for the minimal confidence),}$$

$$\left.\begin{array}{l} s.t. \\ \mu_{\tilde{x}_1}(x_1)=5(x_1-0.6)-5(|x_1-0.8|+x_1-0.8), \\ \mu_{\tilde{x}_2}(x_2)=10(x_2-0.1)-10(|x_2-0.2|+x_2-0.2), \\ \mu_{\tilde{x}_3}(x_3)=10(x_3-0.1)-15(|x_3-0.2|+x_3-0.2), \\ \mu_{\tilde{x}_4}(x_4)=25(x_4-0.01)-17.5(|x_4-0.05|+x_4-0.05), \\ \mu_{\tilde{x}_5}(x_5)=5(x_5-0.6)-5(|x_5-0.8|+x_5-0.8), \\ \mu_{\tilde{x}_6}(x_6)=10(x_6-0.7)-10(|x_6-0.8|+x_6-0.8), \\ \mu_{\tilde{x}_7}(x_7)=10(x_7-0.7)-5(|x_7-0.8|+x_7-0.8)-5(|x_7-0.85|+x_7-0.85), \\ \mu_{\tilde{x}_8}(x_8)=25(x_8-0.01)-12.5(|x_8-0.05|+x_8-0.05) \\ \qquad\qquad -25(|x_8-0.07|+x_8-0.07), \end{array}\right\} \quad (14)$$

$$\left.\begin{array}{l} \alpha[0.2x_1x_3(1\text{-}\tilde{x}_5)0.8+0.2(1\text{-}\tilde{x}_1)\tilde{x}_3(1\text{-}\tilde{x}_6)0.8 \\ +0.2\tilde{x}_1(1-\tilde{x}_3)(1-\tilde{x}_7)0.6+0.2(1-\tilde{x}_1)(1-\tilde{x}_3)(1-\tilde{x}_8)0.6 \\ +0.8\tilde{x}_2\tilde{x}_4(1-\tilde{x}_5)0.8+0.8(1-\tilde{x}_2)\tilde{x}_4(1-\tilde{x}_6)0.8 \\ +0.8\tilde{x}_2(1-\tilde{x}_4)(1-\tilde{x}_7)0.6+0.8(1-\tilde{x}_2)(1-\tilde{x}_4)(1-\tilde{x}_8)0.6]=1, \\ 0.1\le P(+a\,|\,\text{-}d,+e)\le 0.3, \\ P(+b\,|\,\text{-}d,+e)\le 2P(+c\,|\,\text{-}d,+e), \\ Either\quad P(+a\,|\,\text{-}d,+e)\le 0.2\quad or\quad P(+b\,|\,\text{-}d,+e)\le 0.2, \\ \tilde{x}_i\in F(feasible\quad set). \end{array}\right\} \quad (15)$$

First (14) is linearized using Proposition 2 and then the initial program is altered into the equivalent program as follows.

$$Maximize\quad \mu_{\tilde{x}_i}(x_i),\ \ i=1,2,\cdots,8,\ \ \text{(for the maximal confidence), or}$$

$$Minimize\quad \mu_{\tilde{x}_i}(x_i),\ \ i=1,2,\cdots,8,\ \ \text{(for the minimal confidence),}$$

s.t.

$$\left.\begin{aligned}
&\mu_{\tilde{x}_1}(x_1) = 5(x_1 - 0.6) - 2[5(x_1 - 0.8 + d_1)],\\
&\mu_{\tilde{x}_2}(x_2) = 10(x_2 - 0.1) - 2[10(x_2 - 0.2 + d_2)],\\
&\mu_{\tilde{x}_3}(x_3) = 10(x_3 - 0.1) - 2[15(x_3 - 0.2 + d_3)],\\
&\mu_{\tilde{x}_4}(x_4) = 25(x_4 - 0.01) - 2[17.5(x_4 - 0.05 + d_4)],\\
&\mu_{\tilde{x}_5}(x_5) = 5(x_5 - 0.6) - 2[5(x_5 - 0.8 + d_5)],\\
&\mu_{\tilde{x}_6}(x_6) = 10(x_6 - 0.7) - 2[10(x_6 - 0.8 + d_6)],\\
&\mu_{\tilde{x}_7}(x_7) = 10(x_7 - 0.7) - 2[5(x_7 - 0.8 + d_{71}) + 5(x_7 - 0.85 + d_{71} + d_{72})],\\
&\mu_{\tilde{x}_8}(x_8) = 25(x_8 - 0.01) - 2[12.5(x_8 - 0.05 + d_{81}) + 25(x_{82} - 0.07 + d_{81} + d_{82})],\\
&x_1 + d_1 \geq 0.8,\ \ 0 \leq d_1 \leq 0.8,\\
&x_2 + d_2 \geq 0.2,\ \ 0 \leq d_2 \leq 0.2,\\
&x_3 + d_3 \geq 0.2,\ \ 0 \leq d_3 \leq 0.2,\\
&x_4 + d_4 \geq 0.05,\ \ 0 \leq d_4 \leq 0.05,\\
&x_5 + d_5 \geq 0.8,\ \ 0 \leq d_5 \leq 0.8,\\
&x_6 + d_6 \geq 0.8,\ \ 0 \leq d_6 \leq 0.8,\\
&x_7 + d_{71} + d_{72} \geq 0.85,\ \ 0 \leq d_{71} \leq 0.8,,\ \ 0 \leq d_{72} \leq 0.05,\\
&x_8 + d_{81} + d_{82} \geq 0.07,\ \ 0 \leq d_{81} \leq 0.05,,\ \ 0 \leq d_{82} \leq 0.02, \quad \textit{and} \quad (15)
\end{aligned}\right\} \qquad (16)$$

To ensure belief propagation the lower bound of the membership functions is set at 0.2; that is, the membership of every fuzzy parameter must equal or exceed 0.2, which excludes scenarios involving poorly estimated parameters.

LINGO 8.0 solves Example 2.1 in less than one second. The solutions for maximal confidence are α =2.6743 and

$$P(+a \mid +d, -e) = \alpha P(+a) \sum_{i=1}^{4} P(z_i \mid +a) P(-d \mid z_i) P(+e \mid z_i) = 0.1097,$$

$$P(+b \mid +d, -e) = \alpha \sum_{a=0,1} [P(a) \sum_{i=1,3} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] = 0.20,$$

$$P(+c \mid +d, -e) = \sum_{a=0,1}^{+a} [P(a) \sum_{i=1,2} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] = 0.1.$$

The solutions for minimal confidence are

$$P(+a \mid +d, -e) = \alpha P(+a) \sum_{i=1}^{4} P(z_i \mid +a) P(-d \mid z_i) P(+e \mid z_i) = 0.1056,$$

$$P(+b \mid +d, -e) = \alpha \sum_{a=0,1} [P(a) \sum_{i=1,3} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] = 0.2,$$

$$P(+c \mid +d, -e) = \sum_{a=0,1}^{+a} [P(a) \sum_{i=1,2} P(z_i \mid a) P(-d \mid z_i) P(+e \mid z_i)] = 0.1 \quad \square$$

Table 3 lists the detailed solutions.

The results of this model differ from those for Example 1. In Table 3, $P(+a \mid +d, -e)$ changes to [0.1056, 0.1058], where 0.1056 and 0.1058 is solved by minimizing and maximizing the fuzzy membership functions, respectively. $P(+b \mid +d, -e)$ changes to 0.2, and $P(+b \mid +d, -e)$ changes to 0.1, which implies that the solutions are insensitive to the confidence of fuzzy parameters. This variance results from the constraints that dominate the belief propagation. Readers may have deduced that Example 1 can be considered a special case in which every membership of the fuzzy parameters converges on 1.

Table 3: Solution of Example 2.1

	Under maximal confidence	Under minimal confidence
$BEL(a+)$	0.1058	0.1056
$BEL(b+)$	0.20	0.2
$BEL(c+)$	0.10	0.10
x_1	0.9449	0.96
x_2	0.2762	0.28
x_3	0.2339	0.2062
x_4	0.1284	0.1300
x_5	0.6458	0.64
x_6	0.7265	0.72
x_7	0.7321	0.7405
x_8	0.0831	0.0860
$\mu_{\tilde{x}_1}(x_1)$	0.2755	0.2
$\mu_{\tilde{x}_2}(x_2)$	0.2383	0.2
$\mu_{\tilde{x}_3}(x_3)$	0.3218	0.2
$\mu_{\tilde{x}_4}(x_4)$	0.2162	0.2
$\mu_{\tilde{x}_5}(x_5)$	0.2290	0.2
$\mu_{\tilde{x}_6}(x_6)$	0.2651	0.2
$\mu_{\tilde{x}_7}(x_7)$	0.2496	0.2
$\mu_{\tilde{x}_8}(x_8)$	0.3441	0.2

Under certain circumstances, knowledge workers may need to compromise among diverse, even conflicting information sources, causing fuzzy parameters to differ from their most possible values.

<u>Example 2.2</u> (Just-in-time techniques and firm performance): This example uses the Bayesian network to model the relationship between just-in-time purchasing techniques and firm performance [10]. Just-in-time purchasing (JITP) is an important component of supply chain management in managing inventory flows. Several key factors link the JITP process and firm performance, and Figure 6 models the relationships among these factors. Tables 4 and 5 summarize the probability distributions of the nodes and fuzzy parameters.

This study hypothesizes a scenario in which inventory management performance is good ($im+$), employ relationship is poor ($er-$), transportation performance is good ($ta+$), and financial and market performance is poor ($fm-$). The problem involves calculating the belief distribution of all unknown nodes, top management commitment (tp), supplier value-added (su), training (tr), quantity delivered (qd), and time-based quality performance (tq). The reasoning model is formulated as (17).

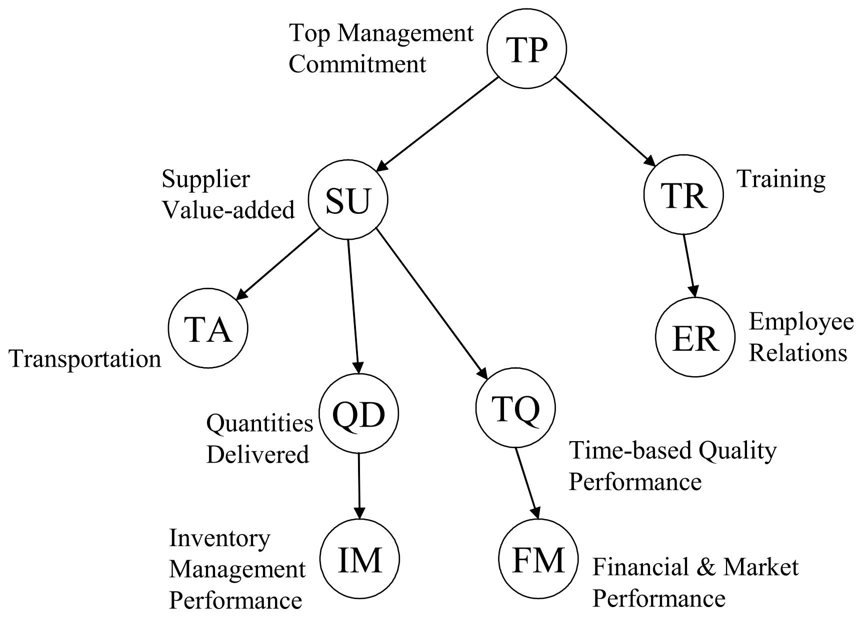

Figure 6: A Bayesian network of the relationships between JITP techniques and performance measures [10]

Table 4: The conditional probability distribution of Example 2.2

$P(tp+) = \tilde{x}_{31}$	
$P(su+ \mid tp+) = \tilde{x}_{32}$	$P(su+ \mid tp-) = \tilde{x}_{33}$
$P(tr+ \mid tp+) = \tilde{x}_{34}$	$P(tr+ \mid tp-) = \tilde{x}_{35}$
$P(ta+ \mid su+) = 0.7$	$P(ta+ \mid su-) = 0.1$
$P(qd+ \mid su+) = 0.8$	$P(qd+ \mid su-) = 0.3$
$P(im+ \mid qd+) = 0.3$	$P(im+ \mid qd-) = 0.1$
$P(tq+ \mid su+) = 0.4$	$P(tq+ \mid su-) = 0.05$
$P(fm+ \mid tq+) = 0.7$	$P(fm+ \mid tq-) = 0.1$
$P(er+ \mid tr+) = 0.6$	$P(er+ \mid tr-) = 0.1$

Table 5: The membership functions of fuzzy probabilities in Example 2.2.

Parameter	$\mu_{\tilde{x}_{3i}}(x_{3i})$	Domain of x_{3i}
$\tilde{x}_{31}$	$5(x_{31}-0.1)-7.5(\lvert x_{31}-0.3\rvert+x_{31}-0.3)$	[0.1,0.4]
$\tilde{x}_{32}$	$5(x_{32}-0.4)-5(\lvert x_{32}-0.6\rvert+x_{32}-0.6)$	[0.4,0.8]
$\tilde{x}_{33}$	$20(x_{33}-0.05)-20(\lvert x_{33}-0.75\rvert+x_{33}-0.1)$	[0.05,0.15]
$\tilde{x}_{34}$	$10(x_{34}-0.5)-5(\lvert x_{34}-0.6\rvert+x_{34}-0.6)-5(\lvert x_{34}-0.7\rvert+x_{34}-0.7)$	[0.5,0.8]
$\tilde{x}_{35}$	$10(x_{35}-0.1)-5(\lvert x_{35}-0.2\rvert+x_{35}-0.2)-5(\lvert x_{35}-0.3\rvert+x_{35}-0.3)$	[0.1,0.4]

Maximize $\mu_{\tilde{x}_{3i}}(x_{3i})$ (for the maximal confidence), or

Minimize $\mu_{\tilde{x}_{3i}}(x_{3i})$ (for the minimal confidence),

s.t.

$$
\begin{aligned}
&\mu_{\tilde{x}_{31}}(\tilde{x}_{31})=5(x_{31}-0.1)-7.5(\lvert x_{31}-0.3\rvert+x_{31}-0.3),\\
&\mu_{\tilde{x}_{32}}(x_{32})=5(x_{32}-0.4)-5(\lvert x_{32}-0.6\rvert+x_{32}-0.6),\\
&\mu_{\tilde{x}_{33}}(x_{33})=20(x_{33}-0.05)-20(\lvert x_{33}-0.75\rvert+x_{33}-0.1),\\
&\mu_{\tilde{x}_{34}}(x_{34})=10(x_{34}-0.5)-5(\lvert x_{34}-0.6\rvert+x_{34}-0.6)-5(\lvert x_{34}-0.7\rvert+x_{34}-0.7),\\
&\mu_{\tilde{x}_{35}}(x_{35})=10(x_{35}-0.1)-5(\lvert x_{35}-0.2\rvert+x_{35}-0.2)-5(\lvert x_{35}-0.3\rvert+x_{35}-0.3),\\
&\alpha\sum_{tp}\sum_{su}\sum_{tr}\sum_{qd}\sum_{tq}[P(tp)P(tr\mid tp)P(su\mid tp)P(ta+\mid su)P(qd\mid su)P(tq\mid su)\\
&\times P(er-\mid tr)P(im+\mid qd)P(fm-\mid tq)]=1,\\
&P(tp+\mid ta+,er-,im+,fm-)>0.6,\\
&P(su+\mid ta+,er-,im+,fm-)>0.8.
\end{aligned}
\tag{17}
$$

First the nonlinear membership functions are linearized, yielding (18).

Maximize $\mu_{\tilde{x}_{3i}}(x_{3i})$ (for the maximal confidence), or

Minimize $\mu_{\tilde{x}_{3i}}(x_{3i})$ (for the minimal confidence),

$$
\begin{aligned}
& s.t. \\
& \mu_{\tilde{x}_{31}}(x_{31}) = 5(x_{31}-0.1) - 2[7.5(x_{31}-0.3+d_{31})], \\
& \mu_{\tilde{x}_{32}}(x_{32}) = 5(x_{32}-0.4) - 2[7.5(x_{32}-0.6+d_{32})], \\
& \mu_{\tilde{x}_{33}}(x_{33}) = 20(x_{33}-0.05) - 2[20(x_{33}-0.1+d_{33})], \\
& \mu_{\tilde{x}_{34}}(x_{34}) = 10(x_{34}-0.5) - 2[5(x_{34}-0.6+d_{341}) + 5(x_{34}-0.7+d_{342})], \\
& \mu_{\tilde{x}_{35}}(x_{35}) = 10(x_{35}-0.1) - 2[5(x_{35}-0.2+d_{351}) + 5(x_{35}-0.3+d_{352})], \\
& \alpha \sum_{tp}\sum_{su}\sum_{tr}\sum_{qd}\sum_{tq}[P(tp)P(tr \mid tp)P(su \mid tp)P(ta+ \mid su)P(qd \mid su)P(tq \mid su) \\
& \times P(er- \mid tr)P(im+ \mid qd)P(fm- \mid tq)] = 1, \\
& P(tp+ \mid ta+, er-, im+, fm-) > 0.6, \\
& P(su+ \mid ta+, er-, im+, fm-) > 0.8.
\end{aligned} \tag{18}
$$

LINGO 8.0 solves the above program in approximately 5 seconds, obtaining the following results

For the model under the maximal confidence:

α =30.5359,

$P(tp+ \mid ta+, er-, im+, fm-) = 0.6103,$
$P(su+ \mid ta+, er-, im+, fm-) = 0.8,$
$P(tr+ \mid ta+, er-, im+, fm-) = 0.2886,$
$P(qd+ \mid ta+, er-, im+, fm-) = 0.8510,$
$P(tq+ \mid ta+, er-, im+, fm-) = 0.1489.$

For the model under the minimal confidence:

α =30.9791,

$P(tp+ \mid ta+, er-, im+, fm-) = 0.6000,$
$P(su+ \mid ta+, er-, im+, fm-) = 0.8,$
$P(tr+ \mid ta+, er-, im+, fm-) = 0.3695,$
$P(qd+ \mid ta+, er-, im+, fm-) = 0.8510,$
$P(tq+ \mid ta+, er-, im+, fm-) = 0.1489.$

Table 6 lists the details.

3.5 Discussions and conclusions

This chapter develops a non-linear programming model for dealing with constrained abductive reasoning on Bayesian networks. This model can be built on any exact propagation methods in Bayesian networks. The present study involves some fuzzy parameters and certain extra constraints. Optimization techniques, including piecewise linearization, are adopted to solve this non-linear programming model and obtain the solutions to the abductive reasoning problems under maximal and minimal confidence to the fuzzy parameters. Since the constraints in this model are extremely non-linear, and numerous non-separable terms are involved, local optima are obtained at the present stage. To enhance the solution quality, some global optimization techniques [24,40,41] can be further used for extended studies. Simultaneously, various reasoning related constraints are considered, including boundary constraints, dependency and disjunctive constraints. Compared to traditional methods that deal with constraints by dummy auxiliary nodes [8, 10], this optimization model of abduction avoids network restructuring. All extra information related to reasoning is considered to be additional constraints in the proposed non-linear program.

Table 6: Solution of Example 2.2

	Under maximal confidence	**Under minimal confidence**
$BEL(tp+)$	0.6103	0.6000
$BEL(su+)$	0.8000	0.8000
$BEL(tr+)$	0.2886	0.3694
$BEL(qd+)$	0.8510	0.8510
$BEL(tq+)$	0.1489	0.1489
x_{31}	0.3567	0.3561
x_{32}	0.7098	0.8
x_{33}	0.1274	0.1207
x_{34}	0.5451	0.7
x_{35}	0.3549	0.3
$\mu_{\tilde{x}_{31}}(x_{31})$	0.4329	0
$\mu_{\tilde{x}_{32}}(x_{32})$	0.4510	0
$\mu_{\tilde{x}_{33}}(x_{33})$	0.4510	0
$\mu_{\tilde{x}_{34}}(x_{34})$	0.4510	0
$\mu_{\tilde{x}_{35}}(x_{35})$	0.4510	0

Chapter 4 Diagnosis and decision with fuzzy parameters

This chapter discusses the reasoning systems that need to complete diagnosis and decision making simultaneously. In this class of problems, the knowledge base will be extended into an influence diagram, in which the decision variables and fuzzy parameters are introduced. The problem of this chapter is presented as follow.

Problem 3: Given the evidence set $\check{E}$ from $BN_3 = (DN, XN, L, \tilde{P})$, compute the belief distribution of $\hat{U} \subset BN_3 \backslash \check{E}$, $BEL(\hat{U} | \check{E})$ ■

In some environments, such as in a medical reasoning system, two generic reasoning tasks are vital: diagnostic reasoning and treatment planning. Diagnostic reasoning is the process of reconstructing the past facts from the observed evidence. Treatment planning is reasoning about the effects of actions treated on patients [27]. Usually, the practices of medicine and business require both kinds of reasoning to work simultaneously. However, few current reasoning methods can conduct the two reasoning tasks successfully at one time. Besides, the reasoning systems become more complex considering the complexity of human bodies and its relationships with the regional factors.

In some clinical cases, various factors may raise the difficulty in reasoning, such as the demographic variances of nosography, the incomplete knowledge of the diseases (e.g. Severe Acute Respiratory Syndrome, SARS, in the early 2003), some restrictions on estimating relevant parameters of the diseases, etc. In these cases, the clinicians' experiences and judgment may be very useful to diagnosis and prescription. Therefore, the site-by-site factors and clinicians' knowledge, which may be expressed with extra constraints in the reasoning systems, need to be integrated into the medical decision support systems. At the same time, owing to the difficulties to estimate the causal effects between possible pathogens and the diseases, the parameters of the knowledge base can be expressed as fuzzy numbers.

Considering the clinical issues mentioned above, the authors are motivated to develop a methodology with the following features.

1. Complete diagnostic reasoning as well as treatment planning.
2. Combine the formal knowledge base as well as decision-makers' judgments that present as extra constraints.
3. Work compatibly with the circumstance where fuzzy information is involved.

In the following section, the background of this research and the proposed approach will be interpreted.

4.1 Influence diagrams

In medical informatics and industrial domains, Bayesian networks and influence diagrams [30,31,33,35,39] are widely used knowledge representation and decision aids under uncertainty. Influence diagrams are directed acyclic graphs with three types of nodes: decision nodes, chance nodes, and a value node. Decision nodes, shown as squares, represent choices available to the decision-makers. Chance nodes, shown as circles, represent random variables (or uncertain quantities). Finally, the value node, shown as a diamond, represents the objective (or utility) to be maximized. In a multiple objective decision making model, there may be more than one value nodes.

However, two limitations still persist when utilizing the above approaches for solving medical reasoning problems:

1. All associated probabilities are assumed to be crisp values.
2. Difficult to introduce the constraint among the nodes in Bayesian networks or influence diagrams.

3. Planning and diagnostic problems are not considered in one paradigm.

The limitations mentioned above restrict the practical usefulness of medical reasoning on Bayesian networks and influence diagrams in the following facts. First, the conditional probabilities between a node and its parent nodes could be fuzzy instead of a crisp numbers, owing to the difficulties of learning accurately the cause-effect relationships among the nodes. Second, as a common fact, the experts may have some professional speculations in the form of constraints between the nodes in a Bayesian network. These constraints could be boundary, dependency, or disjunctive conditions. Third, the investigators of influence diagrams used to maximize the utility functions by node removal processes [30,33,39] and ignore diagnostic reasoning tasks; on the other hand, Bayesian networks have been used widely in probabilistic reasoning but lacked the capability to suggest the optimal decision.

This section proposes an optimization model to make diagnostic reasoning and treatment planning for bacterial infections, where the cause-effect relationships are expressed with an influence diagram and fuzzy data. The inputs of the reasoning system are conditional probability distributions of the network nodes, the associated costs of the candidate antibiotic treatments, the expected effects of the treatments, and extra constraints regarding belief propagation. Since the prevalence of the pathogens and infections are determined by many site-by-site factors and subjective knowledge, the decision may involve uncertainty not compliant with conventional approaches and quite different background. So we allow the decisions to be made under fuzzy environments, at which some of the parameters could be fuzzy parameters [7], and some constraints regarding diagnosis are introduced. When a patient is received, this reasoning system can, based on the present symptoms or bacteriological tests, help the clinician make precise diagnosis at the first decision point, and also supply the suggestions of optimal treatment for the infection. The outputs of the reasoning model are the likelihood of a bacterial infection, the most likely pathogen(s), the suggestion of optimal treatment, the gain of life expectancy of the patient related to the optimal treatment, the probability of coverage from an infection associated with the antibiotic treatment, and the cost-effect analysis of the treatment prescribed. The input-process-output diagram is depicted in Figure 7.

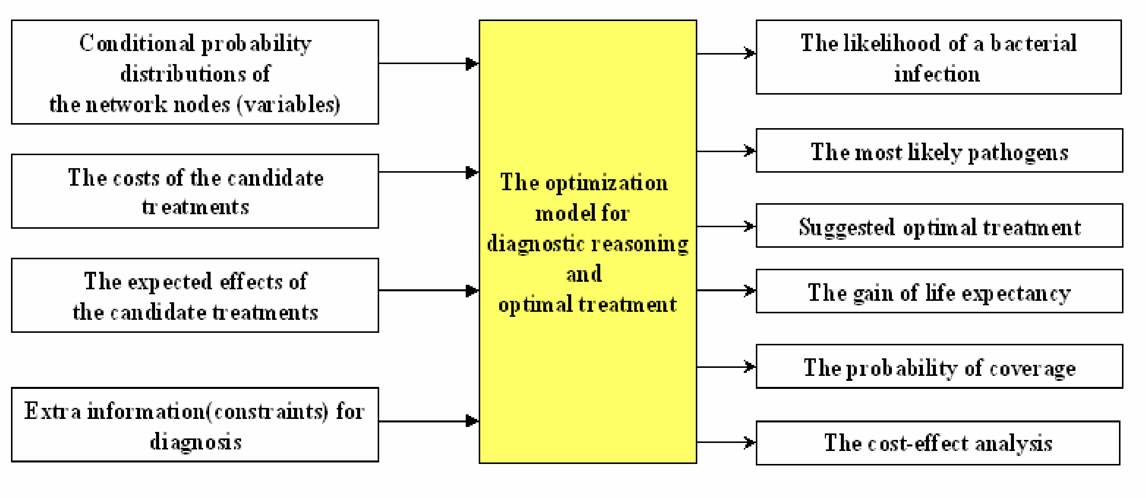

Figure 7: The input-output diagram of the optimization model for Chapter 4

In the following, the authors will introduce an example of Urinary tract infection (UTI), the problem and design goal, and handling the fuzzy information sequentially.

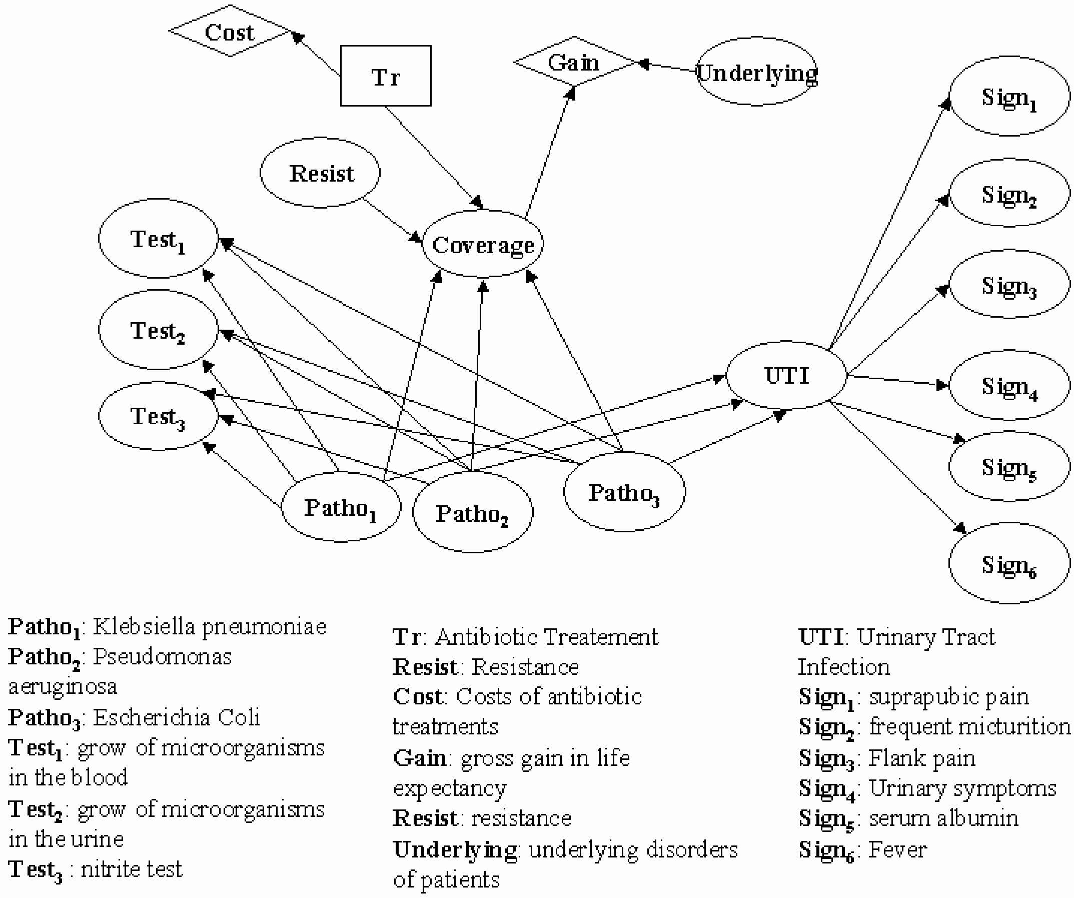

Figure 8: A revised Bayesian network for Urinary tract infection [23]**

Consider one example of urinary tract infections modified from Leibovici et al [23]. As depicted in Figure 8, this example uses an influence diagram as the knowledge and decision model where the conditional probability distributions for the relevant random and decision variables are calculated. For the sake of simplicity and without loss of generality, all random nodes are assumed and organized as binary. The conditional probability distributions of the variables are given as an example in Table 7 through Table 9. The nodes and their states in Figure 8 are described as follow.

** In the latter part of Figure8, the authors put pairs of (**Node_Name**: Description) for each node in the network to explain what the nodes represent.

Table 7: The conditional probabilities of pathogens, tests, and signs of UTI

$P(+patho_1)$=0.1
$P(+patho_2)$=0.09
$P(+patho_3)$=0.09
$P(+uti\,\vert\,+patho_1,+patho_2,+patho_3)=\tilde{x}_1$ $P(+uti\,\vert\,+patho_1,-patho_2,+patho_3)=\tilde{x}_2$ $P(+uti\,\vert\,+patho_1,+patho_2,-patho_3)=\tilde{x}_3$ $P(+uti\,\vert\,+patho_1,-patho_2,-patho_3)=\tilde{x}_4$ $P(+uti\,\vert\,-patho_1,+patho_2,+patho_3)=\tilde{x}_5$ $P(+uti\,\vert\,-patho_1,-patho_2,+patho_3)=\tilde{x}_6$ $P(+uti\,\vert\,-patho_1,+patho_2,-patho_3)=\tilde{x}_7$ $P(+uti\,\vert\,-patho_1,-patho_2,-patho_3)=\tilde{x}_8$

Pathogen ($Patho_i$): a microorganism capable of causing urinary tract infection. For the convenience of computation, only 3 of 12 pathogens are presented: $Patho_1$ (Klebsiella pneumoniae) $Patho_2$ (Pseudomonas aeruginosa), $Patho_3$ (Escherichia Coli). The states of this kind of nodes are severity: severe ($Patho_i$=1) and not severe ($Patho_i$=0).

Urinary tract infection (UTI): The states of this node are severe (UTI=1) and not severe (UTI=0).

Signs and symptoms of urinary tract infection ($Sign_i$): the manifestations that might cause from UTI. There are six possible signs presented in Figure 7: $Sign_1$ (suprapubic pain), $Sign_2$ (Frequent micturition), $Sign_3$ (Flank pain), $Sign_4$ (Urinary symptoms), $Sign_5$ (Serum albumin) and $Sign_6$ (Fever). The states of these nodes are present ($Sign_i$=1) and absent ($Sign_i$=0).

Table 8: The conditional probabilities of Signs ($Sign_i$)

$P(+sign_1 \mid +uti) = 0.6$	$P(+sign_1 \mid -uti) = 0.01$
$P(+sign_2 \mid +uti) = 0.9$	$P(+sign_2 \mid -uti) = 0.10$
$P(+sign_3 \mid +uti) = 0.6$	$P(+sign_3 \mid -uti) = 0.05$
$P(+sign_4 \mid +uti) = 0.8$	$P(+sign_4 \mid -uti) = 0.05$
$P(+sign_5 \mid +uti) = 0.6$	$P(+sign_5 \mid -uti) = 0.10$
$P(+sign_6 \mid +uti) = 0.7$	$P(+sign_6 \mid -uti) = 0.01$

Table 9: Conditional probabilities of Coverage with resistance ($Resist$ =1)

	The instance of ($Patho_1, Patho_2, Patho_3$)							
Treatment*	(1,1,1)	(1,0,1)	(1,1,0)	(1,0,0)	(0,1,1)	(0,0,1)	(0,1,0)	(0,0,0)
tr_0**	0.3	0.4	0.4	0.5	0.4	0.3	0.3	0.6
tr_1	0.7	0.9	0.99	0.95	0.7	0.8	0.75	0.7
tr_2	0.7	0.7	0.85	0.7	0.85	0.8	0.99	0.8
tr_3	0.8	0.8	0.87	0.8	0.95	0.99	0.8	0.9
tr_4	0.7	0.95	0.8	0.9	0.8	0.7	0.9	0.95
tr_5	0.8	0.9	0.85	0.9	0.8	0.9	0.9	0.9

* The costs of the tr_0, tr_1, tr_2, tr_3, tr_4, tr_5 are 5000 (the receiving and process costs), 20000, 25000, 30000, 32000 and 50000 dollars, respectively.

** No treatment.

Bacteriological tests ($Test_i$): $Test_1$ (growth of microorganisms in the blood), $Test_2$ (growth of microorganisms in the urine) and $Test_3$ (nitrite test). The states of these nodes are positive ($Test_i$=1) and negative ($Test_i$=0).

Coverage of UTI ($Coverage$): the percent of pathogens of UIT susceptible to an antibiotic drug. The states of this node are significant ($Coverage$=1) and insignificant ($Coverage$=0).

Resistance to antibiotic drugs ($Resist$): the states of this node are resistant ($Resist$=1) and not resistant ($Resist$=0).

Antibiotic treatment (Tr): the treatment will be appropriate if it matches the in-vitro susceptibility of the pathogens. For simplicity of demonstration, we consider 5 of 26 antibiotic drugs and one additional state for no treatment. Thus, we have 6 alternatives, that is $Tr = \{tr_0, tr_1, tr_2, tr_3, tr_4, tr_5\}$, where tr_0 stands for no treatment and tr_i=0 or 1. When tr_i=1, tr_i is prescribed; opposed, tr_i=0 means that tr_i is not prescribed. For the efficiency of computation, we allow only one antibiotic drug at one time, which let it possible to formulate this decision problem as a mixed 0-1integer program. If more than one drug are mixed in the therapy, the mixed treatment will be regarded as another treatment. Notably, this node is a decision node that has effects on the coverage from urinary tract infection.

Cost: a utility node associated with antibiotic treatments ($Cost(tr_i)$).

Gain: the gain in life expectancy obtained by prescribing an antibiotic drug ($Gain$), which is a function of the coverage ($Coverage$) and the underlying disorder of the patient ($Underlying$).

Underlying: the underlying disorder of the patient ($Underlying$), which will be represented by an equivalent base years of remaining life for the simplicity of computation.

Each variable above is characterized by crisp or fuzzy probabilities given the state of its parents. For instance, $UTI \in \{1,0\}$ represents the dichotomy between having urinary tract infection and not having one. $+uti$ stands for the assertion $UTI = 1$ or "urinary tract infection is present", and $-uti$ stands for the negation of $+uti$, i.e., $UTI = 0$.

Denote **Y** as the parameter set of the Bayesian network depicted in Figure 7. The joint probability distribution of this network with treatment tr_i can be expressed as (19).

$$P(y)=\prod_{i=1}^{3}P(patho_i)\times P(uti\,|\,patho_1,patho_2,patho_3)\times\prod_{j=1}^{6}P(sign_j\,|\,uti)$$
$$\times\prod_{k=1}^{3}P(test_k\,|\,patho_1,patho_2,patho_3) \qquad (19)$$
$$\times P(\text{cov}erage\,|\,patho_1,patho_2,patho_3,resist,tr_i).$$

4.2 Problem and goals

The UTI problem can be seen as one case from **Problem 3**, which we present as **Example 3** as below.

<u>**Example 3**</u> Refer to the conditional probabilities in Table 7 and Table 8, and the evidence that a patient is suffering from frequent micturition ($Sign_2$=1), flank pain ($Sign_3$=1) and urinary symptoms ($Sign_4$=1), but has not fallen into a suprapubic pain ($Sign_1$=0), serum albumin ($Sign_5$=0) or fever ($Sign_6$=0). Denote the evidence set Ě = {ě}={$Sign_1$=0, $Sign_2$=1, $Sign_3$=1, $Sign_4$=1, $Sign_5$=0, $Sign_6$=0}. We need to solve the following two problems.

1. Compute the belief distribution of every $Patho_1$, $Patho_2$, $Patho_3$ and UTI.
2. Make the suggestion of the optimal treatment based on the information given in Table 9, assuming the patient with resistance to the antibiotic treatments ($Resist$=1).

At the first decision point, the clinician tends to make the diagnosis without biological test results; that is, the task is reasoning on the subgraph omitting the nodes $Test_i$ and simplified as to compute *P(y| ě)* where *ě* stands for an instance of the evidence set **Ě,** and **Y** shrinks as {$Patho_1$, $Patho_2$, $Patho_3$, UTI, $Coverage$}. This is reasonable since all the test nodes are Barron nodes in this diagram [30,33,39]. If the treatment prescribed at the first time doesn't work, then some biological tests would be further considered. Besides, this model would like to provide the suggestions for the optimal treatment that maximizes the gain of life expectancy and minimizes the total associated costs.

Notice that some of the parameters in Table 7 are not crisp but fuzzy numbers. For instance, $P(+uti\,|\,+patho_1,+patho_2,+patho_3)$ is not a crisp but a fuzzy number instead, say $\tilde{x}_1$, where $P(+uti\,|\,+patho_1,+patho_2,+patho_3)=\tilde{x}_1$, and associated with a membership function $\mu_{\tilde{x}_1}(x_1)$ represented as follows.

$$\mu_{\tilde{x}_1}(x_1)=\begin{cases}5(x_1-0.6)-5(|x_1-0.8|+x_1-0.8), & 0.6\le x_1\le 1,\\ 0, & elsewhere,\end{cases} \tag{20}$$

where "$|*|$" is the absolute value of a term *.

The above expression and Figure 8 mean that the domain of x_1 is between 0.6 and 1.0. If x_1=0.8 then $\mu_{\tilde{x}_1}(x_1)$=1, which implies that x_1=0.8 is the most confident situation. If $x_1\le$0.6 or $x_1\ge$1 then $\mu_{\tilde{x}_1}(x_1)$=0, which is least possible to happen. If x_1=0.7, then $\mu_{\tilde{x}_1}(x_1)$=0.5.

Based on Proposition 1, (20) can be expressed into a general form as (9).

Now this study expresses the membership functions of the fuzzy parameters $\mu_{\tilde{x}_i}(x_i)$ in Table 10. The readers may find that all the eight fuzzy parameters are triangular fuzzy numbers. However, the membership functions in Table 10 involve absolute terms, which is not convenient to compute. Since the membership function in (9) is a nonlinear function to be maximized, this study will use Proposition 2 to formulate the optimization for diagnosis and optimal treatment.

Now we are ready to formulate the optimization model for diagnosis and treatment planning. Here we formulate the diagnostic reasoning and treatment planning problems as an optimization model. The objectives involved in this model are described as in next section.

4.3 Model development

The objectives involved in this model are described below.

1. To maximize the sum of every fuzzy membership functions. That is, we will make the suggestions of optimal treatment under the maximal confidence of the fuzzy information.
2. To maximize the gain in life expectancy.
3. To minimize the total costs of the treatments.

The first objective is to maximize the sum of every fuzzy membership functions. The decision of treatment prescription influences the gain and cost of this therapy. In this problem, the clinician has 6 candidate treatments to choose, where no treatment is included. We represent each antibiotic treatment as a binary variable tr_i (including tr_0 standing for no treatment) and the associated cost as $Cost(tr_i)$. The total cost is $\sum_{i=0}^{5} Cost(tr_i)$. The objective functions can be expressed as the following.

$$Max \quad z_1 = \mu_{\tilde{x}_i}(x_i) \tag{21}$$

$$Max \quad z_2 = E(Gain(Coverage, Underlying)) \tag{22}$$

$$Min \quad z_3 = \sum_{i=0}^{5} Cost(tr_i) \tag{23}$$

where "$E(*)$" stands for the expectation of a term *.

Table 10: The membership functions of fuzzy probabilities

Parameter	$\mu_{\tilde{x}_i}(x_i)$	Domain of $\tilde{x}_i$
$\tilde{x}_1$	$5(x_1-0.6)-5(\lvert x_1-0.8\rvert+x_1-0.8)$	[0.6,1]
$\tilde{x}_2$	$10(x_2-0.7)-10(\lvert x_2-0.8\rvert+x_2-0.2)$	[0.7,0.9]
$\tilde{x}_3$	$20(x_3-0.7)-20(\lvert x_3-0.75\rvert+x_3-0.75)$	[0.7,0.8]
$\tilde{x}_4$	$10(x_4-0.5)-10(\lvert x_4-0.7\rvert+x_4-0.7)$	[0.5,0.7]
$\tilde{x}_5$	$10(x_5-0.7)-10(\lvert x_5-0.8\rvert+x_5-0.8)$	[0.7,0.9]
$\tilde{x}_6$	$20(x_6-0.55)-20(\lvert x_6-0.6\rvert+x_6-0.6)$	[0.55,0.65]
$\tilde{x}_7$	$10(x_7-0.4)-10(\lvert x_7-0.5\rvert+x_7-0.5)$	[0.4,0.6]
$\tilde{x}_8$	$100(x_8)-100(\lvert x_8-0.01\rvert+x_8-0.01)$	[0,0.02]

In (22), we express the gain in life expectancy as a function of the expectation of *Coverage* and *Underlying*. We assume that the underlying disorder and health status can be converted to an equivalent base year, in this case, 35 years, and the gain is a multiple of the base year. It assumes that, in this clinical case, a patient has the ideal 35 years gain of life expectancy if the probability to recover from UTI is 1. Since the literatures [23] show that one-year gained in life can be regarded equivalent to $55,000, we re-write (22) as (24) for unit standardization.

$$z_2' = 55000 \times E(Gain(Coverage)) * 35 \qquad (24)$$

Setting that only one treatment can be chosen at one decision point, we can formulate the total cost function as in (23). Notably, the expected gain is a function of the resistance of antibiotic treatment (given $Resist=1$), the pathogens ($Patho_j$), and the treatment prescribed (tr_i). The reader may refer to their relationships in Table 9. Defining tr_i as a 0-1 variable, the expectation of $Coverage$, $E(Coverage)$ can be computed as

$$E(Coverage \mid \text{Resist}=1) = \alpha \sum_i \sum_{patho_1} \sum_{patho_2} \sum_{patho_3} tr_i \times P(\text{cov}erage \mid patho_1, patho_2, patho_3, resist=1, tr_i)) \quad (25),$$

where α is the normalizing constant, which will be explained in next subsection.

In this optimization program, two categories of constraints must to be satisfied: (i) the constraints regarding the Bayes' Theorem, and (ii) the extra constraints regarding belief propagation. This optimization model can be implemented with various exact propagation methods. We do not intend to discuss here the details of reasoning algorithms, but focus our attentions on how to formulate this problem as an optimization model. This optimization model can be based on any exact methods. The interested readers may refer to the literatures [1,2,34,35,36,38].

Now we formulate the first category of constraints as

$$\begin{aligned} &\sum_y P(y) \\ &= \alpha \sum_{patho_1} \sum_{patho_2} \sum_{patho_3} \sum_{uti} \sum_{\text{coverage}} [\prod_{j=1}^{3} P(patho_j) \times P(uti \mid patho_1, patho_2, patho_3) \\ &\quad \times P(sign_1=0 \mid uti) P(sign_2=1 \mid uti) P(sign_3=1 \mid uti) P(sign_4=1 \mid uti) \\ &\quad \times P(sign_5=0 \mid uti) P(sign_6=0 \mid uti) \\ &\quad \times \sum_{i=0}^{5} P(\text{cov}erage \mid patho_1, patho_2, patho_3, resist=1, tr_i)] = 1, \end{aligned} \quad (26)$$

$$\sum_{i=0}^{5} tr_i = 1, \quad tr_i = 1 \quad or \quad 0, \quad (27)$$

where α is the normalizing constant which ensures that the sum of the probabilities of every instance of y is 1. The constraint in (27) regulates the clinician to prescribe only one treatment in the first decision point.

At the same time, for a given Bayesian network, the users or the experts, such as the clinicians, may have some professional speculations about the features of some nodes and the relationships among them, in some specific diagnostic context. These features and relationships can be identified as the following types of constraints [14].

1. Boundary constraints

   Some conditional probabilities may have upper or lower bounds. For instance, a clinician may speculate that the posterior probability of $Patho_3$ given the evidence should be higher than 0.3 but lower that 0.5, which can be expressed as

   $$0.3 \leq P(+patho_3 \mid \breve{e}) \leq 0.5 \tag{28}$$

2. Dependency constraints

   The beliefs of some nodes in a Bayesian network may exist mutually dependent relationships. For example, a clinician may presume that the posterior probability of $Patho_2$ should be some multiple of $Patho_3$ given the evidence. Such a relationship is expressed as

   $$P(+patho_1 \mid \breve{e}) \leq 0.5P(+patho_3 \mid \breve{e}) \tag{29}$$

3. Disjunctive constraints

   Sometimes the disjunctive condition between the nodes may exist. For example, a doctor may estimate that either $P(+patho_2 \mid \breve{e})$ or $P(+patho_1 \mid \breve{e})$ is equal to or less than 0.4, which is expressed as

   $$\text{Either } P(+patho_2 \mid \breve{e}) \leq 0.4 \text{ or } P(+patho_1 \mid \breve{e}) \leq 0.4 \tag{30}$$

Introducing constraints (28) and (30) into this reasoning system, this optimization program becomes

$$\left.\begin{array}{ll} Max & z_1 \\ Max & z_2' \\ Min & z_3 \\ st. & (26)-(28),(30) \end{array}\right\} \quad (31)$$

Since the disjunctive constraint (30) is a nonlinear constraint, we will linearize it by some 0-1 variables as the following.

$$\left.\begin{array}{l} M(\theta_1-1)\leq P(+patho_2\,|\,e)-0.4\leq M\theta_1+M(1-\theta_2), \\ M(\theta_2-1)\leq P(+patho_1\,|\,e)-0.4\leq M\theta_2+M(1-\theta_1) \\ \varepsilon\leq\theta_2+\theta_1\leq 1. \end{array}\right\} \quad (32)$$

where θ_1 and θ_2 are 0-1 variables, M is a relatively large number, and ε is a relatively small positive number.

There are four instances of θ_1 and θ_2. (i) When θ_1= 1, θ_2=1, (32) turns into $0\leq P(+patho_2\,|\,e)-0.4\leq M$ and $0\leq P(+patho_1\,|\,e)-0.4\leq M$, which are inactive constraint; (ii) When $\theta_1=0$, $\theta_2=1$, we get $-M\leq P(+patho_2\,|\,e)-0.4\leq 0$ and $0\leq P(+patho_1\,|\,e)-0.4\leq 2M$, which means that when $P(+patho_1\,|\,e)\geq 0.4$, $P(+patho_2\,|\,e)$ must be less than or equal to 0.4; (iii) When θ_1 = 1, θ_2 =0, the constraints are $0\leq P(+patho_2\,|\,e)-0.4\leq 2M$ and $-M\leq P(+patho_1\,|\,e)-0.4\leq 0$, which implies that when $P(+patho_2\,|\,e)\geq 0.4$, $P(+patho_1\,|\,e)$ must be less than or equal to 0.2; (iv) When θ_1= 0, $\theta_2=0$, the inequalities become $-M\leq P(+patho_2\,|\,e)-0.4\leq M$ and $-M\leq P(+patho_1\,|\,e)-0.4\leq M$, which are inactive constraints. The third inequalities in (32) exclude the combinations when θ_1= 1, θ_2=1 and θ_1= 0, θ_2=0. To summarize, (32) implies that either $P(+patho_2\,|\,e)\leq 0.4$ or $P(+patho_1\,|\,e)\leq 0.4$ must be satisfied.

4.4 Algorithm and solutions

The model formulated in the previous section is a multiobjective program, so we adopt the fuzzy approach proposed by Zimmermann [44] to solve it. Following the steps described below, the model is solved.

Algorithm 1

Step 1: Get the ideal solutions of every objective.

To obtain the ideal solutions, every objective is optimized independently regardless of other objectives. In (31), we maximize z_1, z_2', and minimize z_3 individually to acquire their ideal solutions z_1^*, $z_2'^*$ and z_3^*, respectively. The ideal values are z_1^*=8, z_2^*=1722198, and z_3^*=5000.

Step 2: Get the anti-ideal solution of every objective.

To obtain the anti-ideal solutions, every objective is computed in the opposite way regardless of other objectives. Now, we minimize z_1, z_2', and maximize z_3 to acquire the associated ideal solutions z_1^-, $z_2'^-$ and z_3^-, respectively. The anti-ideal values are z_1^-=4, z_2^-=733764.5, and z_3^-=40000.

Step 3: Define the membership function of every objective by its ideal and anti-ideal solutions.

With the ideal and anti-ideal solutions of every objective, we can define their membership functions as follow.

$$\mu_{z_k}(z_k) = \frac{z_k - z_k^-}{z_k^* - z_k^-} \tag{33}$$

The membership functions evaluate the degree of fulfillment for every objective.

Step 4: Maximize the minimal membership function of the three objectives.

Using Zimmermann's fuzzy approach for multi-objective programs, the model (32) can be converted into (34).

$$
\left.
\begin{array}{ll}
Max & \lambda \\
s.t. & \lambda \le \mu_{z_1}(z_1) = \dfrac{z_1 - z_1^{-}}{z_1^{*} - z_1^{-}} \\
 & \lambda \le \mu_{z_2'}(z_2) = \dfrac{z_2' - z_2'^{-}}{z_2'^{*} - z_2'^{-}} \\
 & \lambda \le \mu_{z_3}(z_3) = \dfrac{z_3 - z_3^{-}}{z_3^{*} - z_3^{-}} \\
 & (26)-(28),(32),
\end{array}
\right\} \tag{34}
$$

where λ is defined as

$$
\lambda = \min_{1,2,3}(\mu_{z_1}(z_1), \mu_{z_2'}(z_2'), \mu_{z_3}(z_3)) .
$$

In (34), this study intends to search for the maximum of the minimum level of fulfillment for all the objective functions. To avoid the poor estimation of the fuzzy parameters and decision quality, we set the strict lower bound of the membership function of every fuzzy parameter at 0.5. Applying the ideal and anti-ideal values computed in **Step 1** and **Step 2**, (34) is specified as (35).

$$
\left.
\begin{array}{ll}
Max & \lambda \\
s.t. & \lambda \le \dfrac{z_1 - 4}{8 - 4} \\
 & \lambda \le \dfrac{z_2' - 733764.5}{1722198 - 733764.5} \\
 & \lambda \le \dfrac{z_3 - 40000}{5000 - 40000} \\
 & (26)-(28),(32),
\end{array}
\right\} \tag{35}
$$

This study will solve (35) with LINGO 8.0 developed by LINDO Systems Inc. [48]. LINGO is a comprehensive tool designed to build and solve linear, nonlinear and integer optimization models. LINGO provides a completely integrated package that includes a powerful language for expressing optimization models, a full featured environment for building and editing problems, and a set of fast built-in solvers.

LINGO 8.0 solves (35) in 1 second and obtains the optimal treatment as tr_1 ($tr_1 = 1$, $tr_0 = tr_2 = tr_3 = tr_4 = tr_5 = 0$), the normalizing constant α =323.6647, the optimal minimal membership of the objectives λ =0.5714, and the likelihood of every pathogens:

$P(+patho_1 \mid e) = 0.4000,\quad P(+patho_2 \mid e) = 0.3111,$
$P(+patho_3 \mid e) = 0.3884,\quad P(+uti \mid e) = 0.9895.$

The suggested optimal treatment results in a probability of 0.8396 to cover from the urinary tract infection, equivalent gain in life expectancy as $1616259, and the total costs in $20000. Besides, the clinician can make the diagnosis and optimal prescription at the first decision point with an overall confidence of the fuzzy parameters at 0.5714. We also find $\tilde{x}_3$, $\tilde{x}_4$, $\tilde{x}_7$, $\tilde{x}_8$ significantly apart from its most possible values. It makes sense that, under this reasoning context, the experts need to make some subjective judgment or compromise between different, even conflicting information sources, which make the fuzzy parameters apart from their most confident values. The detailed solutions are listed in Table 11.

4.5 Discussions and conclusions

During the implementation of the reasoning model, the authors find the strength of the optimization model. First, the reasoning system allows the clinicians to combine their special judgments or experiences as extra constraints, which supplement the incomplete formal knowledge. This is useful for some newly discovered disease or infections, and increase the flexibility and robustness for various clinical settings. Second, the model completes two major tasks in medical informatics: diagnostic reasoning and treatment planning simultaneously, which is important a requirement for clinical decision support systems. Third, LINGO provides a powerful and efficient computation tool for solving the optimization model, especially when the authors adopt some linearizing techniques to transform the highly nonlinear program. Based on the authors' experiences, LINGO performs better in solving linear programs than solving nonlinear programs.

However, the authors also find several potential challenges in developing the proposed reasoning system. First, as the clinical problems grow larger and more complex, it may be a burden for the clinicians to formulate the model. In some diseases, there may be tens or hundreds of nodes in the networks. The clinicians will have difficulties to estimate the parameters or specify the conditions of their diagnosis and prescription. Therefore, the system needs some experts in knowledge engineering or information management to participate in, which consequently increases the costs to implement. Second, as the scales of network grow larger, computing the belief of the unknown nodes will be more complicated and time-consuming. Some special techniques for belief propagation may be considered, such as clustering, joint tree decomposition, stochastic simulation, and so on [1,2,34,35,36,38]. How to integrate these computation methods and the optimization model will be a critical issue in implementing the reasoning system. Third, as network structures become huge, implementing the optimization model with LINGO will be fairly challenging. LINGO provides several interfaces with other applications, such as Visual C++, Visual Java, Visual Basic, etc. The system developers can bundle LINGO's functionality into their applications, or call functions from within the LINGO models that were written in an external programming language [48]. It will facilitate generating the codes for LINGO models and importing the parameters or input data from other applications.

Table 11: Solution of Example 3

λ	0.5714		
$z_1 = \sum_{i=1}^{8} \mu_{\tilde{x}_i}(x_i)$	6.2857		
$z_2 = E(Gain)$	1616259		
$z_3 = \sum_{i=0}^{5} Cost(tr_i)$	20000		
$P(+patho_1 \mid e)$	0.4000		
$P(+patho_2 \mid e)$	0.3111		
$P(+patho_3 \mid e)$	0.3884		
$P(+uti \mid e)$	0.9895		
Optimal treatment	$tr_1 = 1$, $tr_0 = tr_2 = tr_3 = tr_4 = tr_5 = 0$		
$P(+coverage \mid e, tr_5)$	0.8396		
x_1	0.8001	$\mu_{\tilde{x}_1}(x_1)$	0.9985
x_2	0.8001	$\mu_{\tilde{x}_2}(x_2)$	0.9970
x_3	0.7601	$\mu_{\tilde{x}_3}(x_3)$	0.7972
x_4	0.5500	$\mu_{\tilde{x}_4}(x_4)$	0.5000
x_5	0.8001	$\mu_{\tilde{x}_5}(x_5)$	0.9970
x_6	0.6000	$\mu_{\tilde{x}_6}(x_6)$	0.9960
x_7	0.45	$\mu_{\tilde{x}_7}(x_7)$	0.5000
x_8	0.0070	$\mu_{\tilde{x}_8}(x_8)$	0.5000

Chapter 5 Diagnosis and decision with fuzzy nodes

This chapter discusses the reasoning systems that need to complete diagnosis and decision-making in dynamic environments. In this class of problems, the knowledge base will be extended into a dynamic fuzzy influence diagram, in which the decision variables and fuzzy variables (nodes) are introduced. The problem of this chapter is presented as follow.

Problem 4: Given the evidence set Ĕ from BN_4= (*DN, CN, XN, L, P*), compute the belief distribution of Û $\subset$ BN_4\ Ĕ, *BEL*(Û| Ĕ). ■

5.1 Reasoning in supply chain management

In some domain, the reasoning tasks may involve various types/degrees of uncertainties and become more complex than the conventional problems, such as supply chain management.

Supply chain management has been commonly recognized as a key issue of business success, especially for multinational industries and global markets. It has shifted the management paradigm of enterprises. There are several works on the causal relationships in multi-echelon supply chains, such as the dynamics of buyer-supplier relationships [3,10,12,32], the strategic role of the buying firms in structuring supplier-supplier relationships and supply chain effectiveness [21], the relationship between just-in-time purchasing techniques and supply chain performance [9], diagnostic reasoning in supply chain alliance with static Bayesian networks [15,16,18], and so on. However, there are limited works on design of integrated systematic methodologies for supply chain diagnostics and optimal solution. Naim et al [29] developed a methodology, Quick Scan, to conduct a supply chain oriented business diagnostics in the European automotive supply chain values streams. Quick Scan is a systematic methodology to collect and synthesize the data from a supply chain. One of the main outputs of Quick Scan is the cause-and-effect diagram of the supply chain. The advantages of Quick Scan are (1) data collection and integration, (2) identification of causal relationships in a supply chain, and (3) a systematic and integrated view of supply chain diagnostics.

Reviewing the literatures on supply chain diagnostics, we find some interesting issues worth further discussions and extensions. First, the diagnostic methodologies are not designed in the previous works. Second, the strength and uncertainty of the causal relationships in supply chain diagnostics are not quantified. Third, there is not an integrated framework on diagnosis and decision-making for the optimal supply chain solutions. Fourth, fuzzy information is not considered in the previous works. Motivated by the open issues, this study proposes an integrated framework based on a dynamic influence diagram for the supply chain diagnosis and decision-making. The uncertainties involved in supply chains are captured by fuzzy numbers and membership functions, which turns the decision model into a dynamic fuzzy influence diagram. Using the integrated framework, it works to answer the queries such as "What are possibly the causes of the poor schedule adherence of the two-echelon supply chains?" and "How information and communication technologies can contribute to the supply chain collaboration? What is the optimal technical solution to the supply chain treatment?"

The remainder of this chapter will be organized as follow. Section 5.2 first addresses a case of European automotive supply chains. Section 5.3 then gives an introduction of dynamic models with fuzzy parameters. A dynamic fuzzy influence diagram will be developed for the two-echelon supply chain based on the case studied in section 5.2. Section 5.4 designs the algorithms to conduct diagnostic reasoning in the dynamic fuzzy influence diagram. Next, section 5.5 formulates a fuzzy multi-objective nonlinear programming model for the optimal supply chain solution. Finally, section 5.6 gives the discussions and conclusions. The research framework is depicted as Figure 9

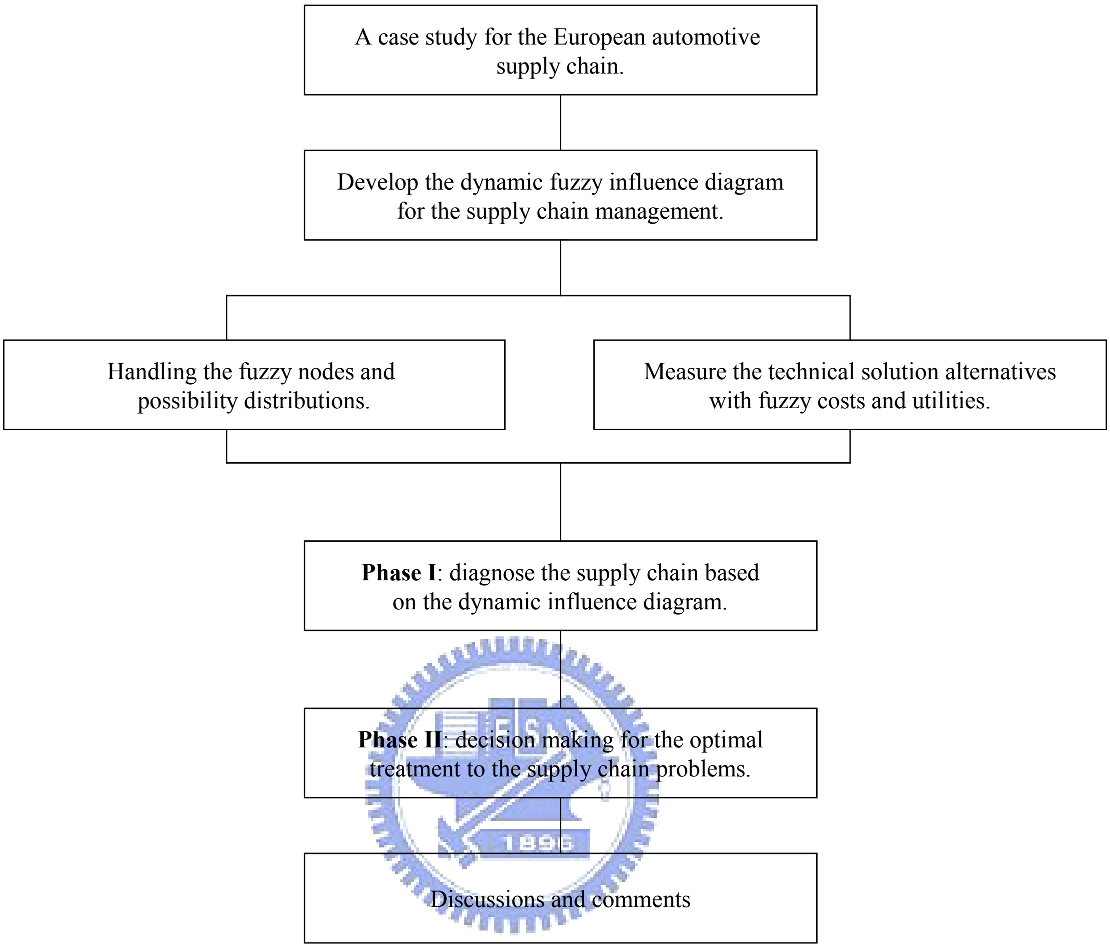

Figure 9: Research framework of Chapter 5

5.2 Problem and Goals

In this section, we first introduce a case of the two-echelon European automotive supply chain [29]. From this case, the readers can realize how the cause and effect diagram of the engine assembler and its component supplier is built and provides a basis for supply chain diagnosis.

Example 4: As one component supplier of the leading engine assembler Company C, Company S puts much attention to its own performance and satisfaction of Company C. From the past periods, Company S observed the poor schedule adherence inside the company, and several help calls from Company C. So Company S initiated an investigation of the company and through the two-echelon supply chain.

Based on Quick Scan, a previous survey to integrate the quantitative and qualitative data from a two-echelon supply chain, all participating companies and third parties who concern the supply chain problems can investigate demand amplification effects in the supply chain and the causes of the poor schedule adherence of the component supplier. One of the main outputs of the Quick Scan is the cause-and-effect diagram shown in Figure 10 [29]. From the cause-and-effect diagram in Figure 10, the readers can find two kinds of factors: the keys of customers (achromatic) and the keys from suppliers (in gray). The arrows in the diagram represent the causal links between the keys of the two-echelon supply chain. Company S traced all the key factors in Figure 10 carefully.

After a thorough investigation, Company S found several symptoms internal his company, which closely relate to some segments in the supply chain. Further collecting relevant data through the supply chain, Company S observed that, for the preceding periods, Company C has put considerable schedule alterations. Company S itself had large product range and high risk of obsolescence, and also kept limited finished goods. The production capacity constraint was high and the scheduling flexibility was low. Besides, the supplier bore large set-up times/ costs and had large batch production.

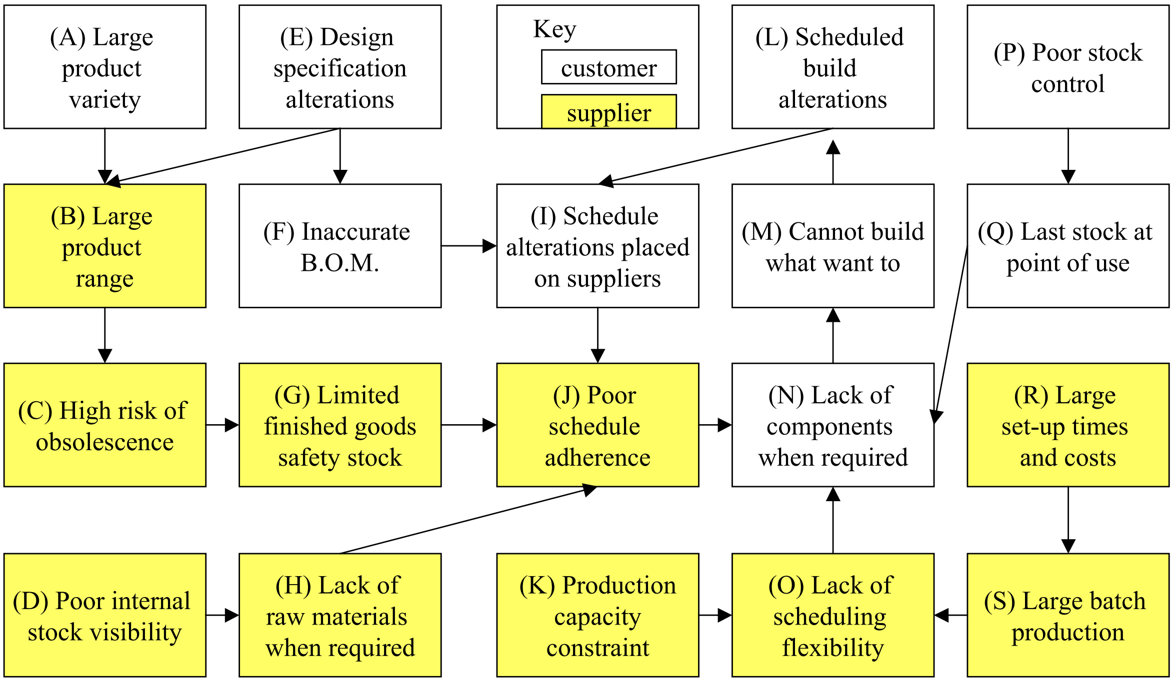

Figure 10: Cause-effect diagram of the two-echelon automotive supply chain [29]

At the same time, Company S foresaw the needs to strengthen the collaborative mechanism in the supply chain, which is supposed to enhance the information transparency, customer satisfaction in the supply chain value stream, and the overall supply chain performance. The information and communication technologies (ICT) are usually believed to facilitate the supply chain collaboration.

In **Example 4**, there are two important issues to concern: (a) What are the causes of poor schedule adherence in the automotive supply chains? How possible are the unknown variables to be the sources of the problems? (b) What is the optimal technical solution to the supply chain collaboration supposed to enhance the supply chain performance?

Since the previous works did not provide proper tools for analyzing the above problems, this study will propose a two-phase model to answer problem (a) and problem (b). Next section will show how the cause-and-effect diagram in Figure 10 can be converted into a dynamic fuzzy influence diagram. Then, diagnostic reasoning in the supply chain will answer the first problem in the diagnostic phase.

5.2.1 Dynamic Bayesian networks with fuzzy nodes

Example 4 can be developed as one application of dynamic Bayesian networks or influence diagrams for supply chain management, involving a leading automotive engine assembler (Company C) and one of its component suppliers (Company S). This study will formulate the supply chain using a dynamic influence diagram.

A static influence diagram can be extended into a dynamic influence diagram [4,8,16,17] by introducing relevant temporal dependencies between representations of the static network at different times. Two types of dependencies can be distinguished in a dynamic network: contemporaneous dependencies and non-contemporaneous dependencies. Contemporaneous dependencies refer to arcs between nodes that represent variables within the same time period. Non-contemporaneous dependencies refer to arcs between nodes that represent variables at different times. We will illustrate how to formulate a dynamic influence diagram for the supply chain diagnosis and treatment, as well as how the participating enterprises in the supply chain can solve the diagnostic problems and obtain the optimal solution on the diagrams.

5.2.2 Uncertainties in supply chains

The supply chain mentioned in Example 4 is a highly uncertain and complex system. The uncertainties originate from several sources.

(a) The stochastic properties in the supply chain system itself, such as the relationships among the nodes in the network, which are random in nature. This category of uncertainty is usually handled with probability theory.

(b) The incomplete knowledge of the system. The decision maker may have only partial knowledge about the system, and need to make some subjective judgments, such as the costs and benefits from the supply chain collaborations. This category of uncertainty is often treated with fuzzy sets and the possibility functions.

(c) The semantic vagueness in the system, such as good manufacturing capability, stock control performance, customer satisfaction, etc. This kind of ambiguity is usually treated with fuzzy sets and the possibility distributions. Once we define the fuzzy nodes and possibility distributions in the influence diagram, it will become a fuzzy influence diagram. Moreover, this study will encounter a dynamic fuzzy influence diagram for the supply chain systems.

This chapter will consider and operate the three types of uncertainty mentioned above.

5.3 Model development

Before applying the graphical decision model to the supply chain problems, the authors first make a brief discussion. According to the uncertainty of the domain problems, there are several types of network nodes: crisp discrete nodes, crisp continuous nodes, fuzzy discrete nodes, and fuzzy continuous nodes. In a word, the crisp nodes behave randomly and are described with a probability distribution. On the other hand, the fuzzy nodes are generally semantically ambiguous or are realized with incomplete knowledge. Considering the system characteristics of the supply chain, we choose crisp discrete and fuzzy continuous nodes to formulate the domain variables. Because there is a feedback loop in Figure 10, the two-echelon supply chain will be expressed by a dynamic fuzzy influence diagram in the following steps. First, this study changes every factor in Figure 10 into a network node. There are two types of nodes: crisp random nodes and fuzzy nodes. For convenience of computation and without loss of generality, each crisp node is assumed two states and turn into a binary variable. The cases of multi-valued crisp nodes are indifferent in nature and intuitive to extend in real-world applications. Each crisp random node is characterized with a prior or conditional probability distribution, which has been learned previously from the field. For the nodes that are semantically ambiguous or incompletely comprehended, we assign each one a possibility distribution to express the level of possibility. The knowledge workers operating on the system can maintain and alter these distributions when any updated information is passed into the systems.

Now Figure 10 is transformed into Figure 11. The description and states of the random nodes are listed in Table 12. Let $\mathbf{X}$ stands for the crisp node set and $\tilde{\mathbf{Y}}$ stands for the fuzzy node set. For the crisp nodes, we use the uppercase letters to represent the variables and lowercase letters for their associated values. For example, $C^t \in \{0,1\}$ represents the dichotomy between low risk of obsolescence and high risk of obsolescence at time t. $+c^t$ stands for C^t =1 and $\neg c^t$ stands for C^t =0. Oppositely, we assign a triangular membership function to represent a fuzzy node. Every fuzzy node is denoted by ($\underline{y}, y^*, \bar{y}$), where $\underline{y}$ represents the left limit, $\bar{y}$ stands for the right limit, and y^* represents the peak of the triangular membership function. This study does not discuss how to learn the relevant parameters, but concentrate on the diagnostic reasoning and decision-making methods. In Figure 11, two kinds of nodes are added: decision node (ICT) and utility node (U). ICT stands for the information and communication technologies that will be selected to improve the supply chain performance and U denotes the utility set.

In Figure 11, we introduce one implicit link from U to I, which is unusual in influence diagrams. The implicit links implies that, the benefit from the collaboration is shown as a value node but will feedback to the next-stage supply chain system and diminish the poor schedule adherence through decreasing schedule alterations on the suppliers.

The diagnostic problems in a supply chain can be regarded from any possible aspects concerning the supply chain. Assume that all the probabilities or possibilities of the nodes have been learned from the historical data and given in Table 13. Remarkably, the probability distributions of some crisp nodes conditioned on the manifestation of their fuzzy parents. To cope with these cases, this study adopts a computable approximate mechanism.

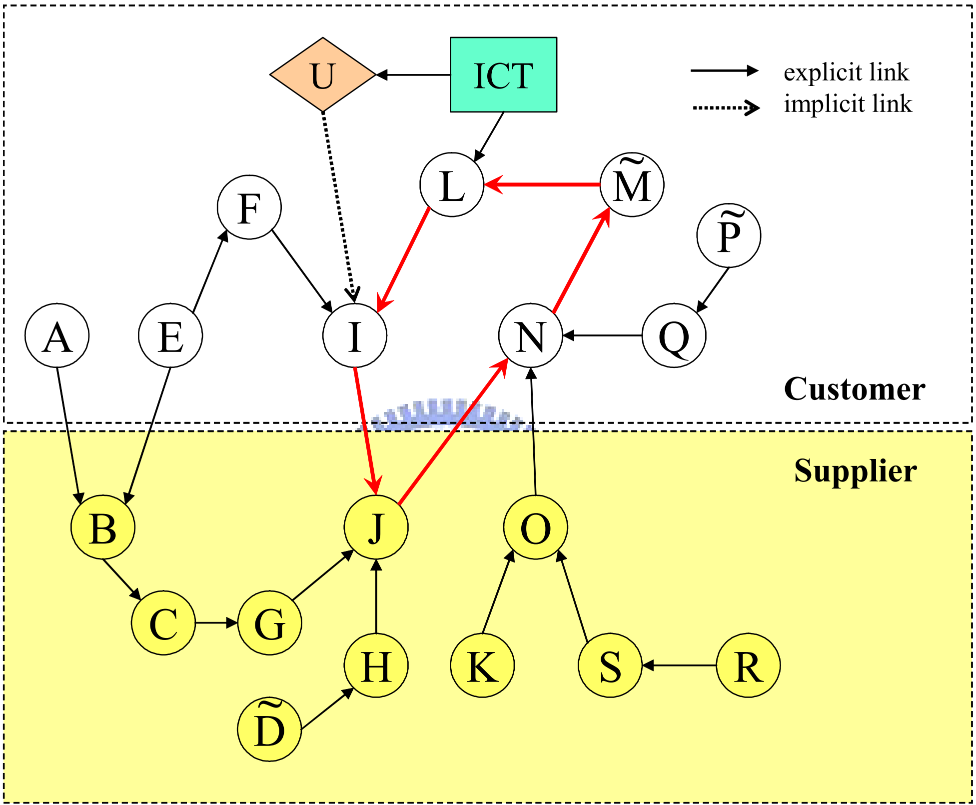

Figure 11: A dynamic influence diagram of the supply chain

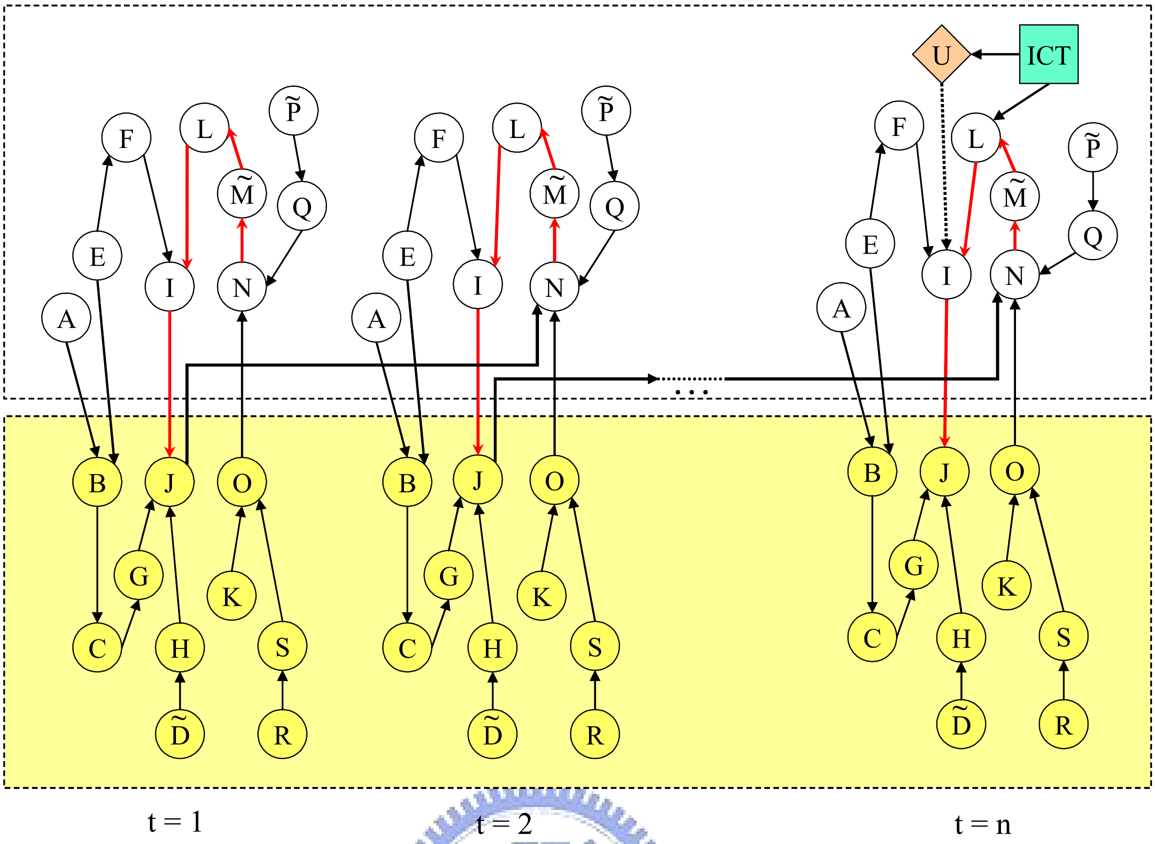

Figure 11a: Time expansion of the dynamic influence diagram in Figure 11

Table 12: Description of nodes in the dynamic influence diagram in Figure 11

Node	Level	Description	State
		Crisp nodes	
A	Customer	Product variety	**1**: > 15 product lines; **0**: <= 15 product lines
B	Supplier	Product range	**1**: > 30 product lines; **0**: <= 30 product lines
C	Supplier	Risk of obsolescence	**1**: > 0.5; 0: <= 0.5
E	Customer	Design specification alterations	**1**: > 2 times/order **0**: <= 2 times /order
F	Customer	B.O.M accuracy	**1**: > 90%, **0**: <= 90%
G	Supplier	Finished goods safety stock	**1**: > 5000 SKUs; **0**: <= 5000 SKUs.
H	Supplier	Lack of raw materials at use	**1**: > 10%; **0**: <= 10%.
I	Customer	Schedule alterations on suppliers	**1**: > 5 times/order **0**: <= 5 times/order
J	Supplier	Schedule adherence	**1**: > 90%; **0**: <=90%
K	Supplier	Production capacity constraint	**1**: > 130% demand orders; **0**: <= 130% demand orders
L	Customer	Schedule build alterations	**1**: > 5 times/order **0**: <= 5 times /order
N	Customer	Lack of components at use	**1**: > 10%; **0**: <=10%
O	Supplier	Scheduling flexibility	**1**: > 35 % capacity; **0**: <= 35 % capacity
Q	Customer	Lost stock at use	**1**: > 5 %; **0**: <= 5 %.
R	Supplier	Set-up times/costs	**1**: >10% standard time **0**: > 10% standard time
S	Supplier	Volume of batch production	**1**: > 60%; **0**: <=60%
		Fuzzy nodes	
$\tilde{D}$	Supplier	Stock control performance	Level of assessed performance
$\tilde{M}$	Customer	Build capability	% of on-time order fulfillment
$\tilde{P}$	Customer	Stock control performance	Level of assessed performance

This work partitions the support of the fuzzy parents into a few sub-domains, and then approximates the crisp children's conditional probability on the sub-domains. In a living expert system, these parameters for the dependency relationships can be estimated and tuned by the knowledge engineers.

The evidence collected till now are poor schedule adherence (J^t=0), considerable schedule alterations (I^t=1), large product range (B^t=1), high risk of obsolescence (C^t=1), limited finished goods (G^t=0), large production capacity constraint (K^t=1), poor scheduling flexibility (O^t=0), large set-up times/ costs (R^t=1) and had large batch production (S^t=1). Given the information on hand, now Company S needs to compute the posterior probability distributions of every proposition in the system backward for n periods, given the evidence set $\check{\mathbf{E}} = \{B^t=1,\ C^t=1,\ G^t=0,\ I^t=1,\ J^t=0,\ K^t=1,\ O^t=0,\ R^t=1,\ S^t=1 | 1 \le t \le n\}$.

In Figure 11, a feedback loops exists among I (schedule alterations placed on suppliers), J (schedule adherence), N (lack of components when required), M (build capability) and L (schedule build alterations). If we take a time expansion aspect, Figure 11 can be expended as Figure 11(a). This study assumes that the relationship and conditional probability distributions among the nodes remain unchanged in the studied horizon. Regardless of the decision and utility nodes temporarily, the joint probability distribution of this dynamic Bayesian network for time t = 1 through n can be expressed as (36).

Table 13: The probability/possibility distributions for the dynamic Influence Diagrams in Example 4

Crisp nodes	
$P(+a^t)$=0.7	
$P(+b^t \mid +a^t,+e^t)$=0.9	$P(+b^t \mid \neg a^t,+e^t)$=0.6
$P(+b^t \mid +a^t,\neg e^t)$=0.8	$P(+b^t \mid \neg a^t,\neg e^t)$=0.2
$P(+c^t \mid +b^t)$=0.85	$P(+c^t \mid \neg b^t)$=0.2
$P(+e^t)$=0.4	
$P(+f^t \mid +e^t)$=0.15	$P(+f^t \mid \neg e^t)$=0.9
$P(+g^t \mid +c^t)$=0.1	$P(+g^t \mid \neg c^t)$=0.8
$P(+h^t \mid d^t_{>0.6})$=0.05	$P(+h^t \mid d^t_{\le 0.6})$=0.9
$P(+i^t \mid +f^t,+l^t)$=0.8	$P(+i^t \mid \neg f^t,+l^t)$=1.0
$P(+i^t \mid +f^t,\neg l^t)$=0.01	$P(+i^t \mid \neg f^t,\neg l^t)$=0.5
$P(+j^t \mid +g^t,+h^t,+i^t)$=0.2	$P(+j^t \mid +g^t,\neg h^t,+i^t)$=0.5
$P(+j^t \mid +g^t,+h^t,\neg i^t)$=0.6	$P(+j^t \mid +g^t,\neg h^t,\neg i^t)$=0.99
$P(+j^t \mid \neg g^t,+h^t,+i^t)$=0.0	$P(+j^t \mid \neg g^t,\neg h^t,+i^t)$=0.5
$P(+j^t \mid \neg g^t,+h^t,\neg i^t)$=0.6	$P(+j^t \mid \neg g^t,\neg h^t,\neg i^t)$=0.8
$P(+k^t)$=0.5	
$P(+l^t \mid m^t_{\ge 0.9})$=0.1	$P(+l^t \mid m^t_{\le 0.8})$=0.9

$P(+n^t \mid +j^{t-1})=0.1$	$P(+n^t \mid \neg j^{t-1})=0.5$
$P(+n^t \mid +o^t, \neg q^t)=0.2$	$P(+n^t \mid \neg o^t, \neg q^t)=0.6$
$P(+n^t \mid +o^t, +q^t)=0.01$	$P(+n^t \mid \neg o^t, +q^t)=0.1$
$P(+o^t \mid +k^t, +s^t)=0$	$P(+o^t \mid \neg k^t, +s^t)=0.7$
$P(+o^t \mid +k^t, \neg s^t)=0.6$	$P(+o^t \mid \neg k^t, \neg s^t)=0.95$
$P(+q^t \mid p^t_{>0.6})=0.1$	$P(+q^t \mid p^t_{\le 0.6})=0.5$
$P(+r^t)=0.5$	
$P(+s^t \mid +r^t)=0.7$	$P(+s^t \mid \neg r^t)=0.3$
Fuzzy nodes	
$Pos(\tilde{d}^t)=(0.3, 0.6, 0.9)$	
$Pos(\tilde{m}^t \mid +n^t)=(0.5,0.6,0.8)$	$Pos(\tilde{m}^t \mid \neg n^t)=(0.9,0.95,1.0)$
$Pos(\tilde{p}^t)=(0.5,0.6,0.7)$	

$$
\begin{aligned}
&L(x,\tilde{y}) = P(x) \otimes Pos(\tilde{y}) \\
&= P(a^1, a^2, \ldots, a^n, b^1, b^2, \ldots, b^n, \ldots, s^1, s^2, \ldots, s^n) \otimes Pos(\tilde{d}, \tilde{m}, \tilde{p}) \\
&= P(a^1)P(b^1 \mid a^1, e^1)P(c^1 \mid b^1)P(e^1)P(f^1 \mid e^1)P(g^1 \mid c^1)P(h^1 \mid \tilde{d}^1) \\
&\times P(i^1 \mid f^1, l^1)P(j^1 \mid i^1, g^1, h^1)P(k^1)P(l^1 \mid \tilde{m}^1)P(n^1 \mid o^1, q^1) \\
&\times P(o^1 \mid s^1, k^1)P(q^1 \mid \tilde{p}^1)P(r^1)P(s^1 \mid r^1) \\
&\otimes [Pos(\tilde{d}^1) \wedge Pos(\tilde{m}^1 \mid n^1) \wedge Pos(\tilde{p}^1)] \\
&\times \prod_{t=2}^{n} [P(a^t)P(b^t \mid a^t, e^t)P(c^t \mid b^t)P(e^t)P(f^t \mid e^t)P(g^t \mid c^t)P(h^t \mid \tilde{d}^t) \\
&\times P(i^t \mid f^t, l^t)P(j^t \mid i^t, g^t, h^t)P(k^t)P(l^t \mid \tilde{m}^t)P(n^t \mid j^{t-1}, o^t, q^t) \\
&\times P(o^t \mid s^t, k^t)P(q^t \mid \tilde{p}^t)P(r^t)P(s^t \mid r^t)] \\
&\otimes [Pos(\tilde{d}^t) \wedge Pos(\tilde{m}^t \mid n^t) \wedge Pos(\tilde{p}^t)].
\end{aligned} \tag{36}
$$

The equation in (36) involves both joint probability and possibility functions, so we use “$\otimes$” to denote the operator link crisp and fuzzy parameters.

The term $P(n^t \mid j^{t-1}, o^t, q^t)$ embraces contemporaneous dependencies at time t and non-contemporaneous dependencies at $t-1$. There are two simple parameterized decompositions used commonly by time-series analysts: the additive and the multiplicative decomposition [4]. The additive decomposition is used commonly in time-series analysis for integrating predictions based on current observations with predictions based on historical observations. Additive decompositions are an integral aspect of models that purport to forecast future values of time-series. The multiplicative decomposition is used usually to model log-linear systems in engineering applications. Both decompositions employ likelihood weights, which provide a language for assigning measures of reliability to information about different periods. Using this approach, we can consider the probabilistic dependencies from contemporaneous sets of variables and from variables at different points in the past as providing independent sources of information. The measures are used to weight the contributions of the contemporaneous and non-contemporaneous dependencies separately. The sum of the predictions, each weighted by its likelihood, gives the final predictions. The use of likelihood weighting allows an expert to specify the weight of the past versus the present easily. Consider the following property.

Additive and multiplicative decomposition

Let ω denotes the likelihood that n^t predicted from the information at period t, and $(1-\omega)$ denotes the likelihood that n^t predicted from the information prior to time t. In the additive decomposition, the conditional probability function $P(n^t \mid j^{t-1}, o^t, q^t)$ can be given by

$$P(n^t \mid j^{t-1}, o^t, q^t) = \omega P(n^t \mid o^t, q^t) + (1-\omega) P(n^t \mid j^{t-1}) \tag{37}$$

In the multiplicative decomposition, the conditional probability function is

$$P(n^t \mid j^{t-1}, o^t, q^t) = \gamma\, P(n^t \mid o^t, q^t)^{\omega} \times P(n^t \mid j^{t-1})^{1-\omega} \qquad (37a)$$

where γ is a constant that normalizes the probability distributions to unify. The likelihood weight ω can be learned from the historical data with maximum likelihood methods [4]. Considering the dynamic properties of Example 4, this study will use additive decomposition in (37).

In modeling the supply chain with an influence diagram, we encounter a dilemma. The two-echelon supply chain is first noticed with its inefficiency in meeting the manufacturing schedule, and needs a first-handed diagnosis prior to a treatment solution. Therefore, this supply chain system needs two things at two different timing: a thorough diagnosis and a suggestion as to the optimal technical solution. Conventional approaches to evaluate the influence diagram do not allow the delay between diagnosis and decision-making. Hence, this study will divide the procedure for treating the supply chain system into two phases: diagnostic phase and optimization phase. In the diagnostic phase, the authors first ignore the decision and utility nodes and regard this system as a Bayesian network. After the origin of supply chain inefficiency is uncovered, next optimization phase will be activated and find the optimal solution.

5.4 Algorithms and solutions

Example 4 is a typical diagnostic reasoning problem in a dynamic environment. We will use the stochastic simulation [2,35] to solve this problem.

5.4.1 Phase I: diagnostic phase

First of all, we denote by w_X the state of all variables except X, then the value of X will be chosen by tossing a coin that favors 1 over 0 by a ratio of $P(+x \mid w_X)$ to $P(\neg x \mid w_X)$. We will show that $P(x \mid w_X)$, the distribution of each crisp variable X conditioned on the values w_X of all other variables in the system, can be calculate by purely local computations. The distributions of $P(x \mid w_X)$ in this network at time t are as (38)-(44).

Distributions of crisp nodes

$$P(a^t \mid w_{A^t}) = \alpha P(a^t) P(b^t \mid a^t, e^t) \tag{38}$$

$$P(e^t \mid w_{E^t}) = \alpha P(e^t) P(b^t \mid a^t, e^t) P(f^t \mid e^t) \tag{39}$$

$$P(f^t \mid w_{F^t}) = \alpha P(f^t \mid e^t) P(i^t \mid f^t, l^t) \tag{40}$$

$$P(h^t \mid w_{H^t}) = \alpha P(h^t \mid \tilde{d}^t) P(j^t \mid g^t, h^t, i^t) \tag{41}$$

$$P(l^t \mid w_{L^t}) = \alpha P(l^t \mid \tilde{m}^t) P(i^t \mid f^t, l^t) \tag{42}$$

$$P(n^t \mid w_{N^t}) = \alpha[\omega P(n^t \mid o^t, q^t) + (1-\omega) P(n^t \mid j^{t-1})] \otimes Pos(\tilde{m}^t \mid n^t) \tag{43}$$

$$P(q^t \mid w_{Q^t}) = \alpha P(q^t \mid p^t)[\omega P(n^t \mid o^t, q^t) + (1-\omega) P(n^t \mid j^{t-1}) \tag{44}$$

Similarly, the distribution of $Pos(\tilde{y} \mid w_{\tilde{Y}})$ at time t are list as (45)-(47).

Distributions of fuzzy nodes

$$Pos(\tilde{d}^t \mid w_{\tilde{D}^t}) = \alpha Pos(\tilde{d}^t) \otimes P(h^t \mid \tilde{d}^t) \tag{45}$$

$$Pos(\tilde{m}^t \mid w_{\tilde{M}^t}) = \alpha Pos(\tilde{m}^t \mid n^t) \otimes P(l^t \mid \tilde{m}^t) \tag{46}$$

$$Pos(p^t \mid w_{\tilde{P}^t}) = \alpha Pos(\tilde{p}^t) \otimes P(q^t \mid \tilde{p}^t) \tag{47}$$

where α is the normalizing constant.

This study assumes ω=0.5 for all periods except the starting period (t=1) when ω

is 1. Assume that the evidence remains unchanged during the simulation. This assumption can be released when applied to more complex scenarios. The algorithm of stochastic simulation is listed as follow.

<u>**Algorithm 2**</u>: **Stochastic simulation for crisp nodes**

Step 1: Read **EvidenceSet**, **UnknownNodeSet**. In Example 4, **EvidenceSet=**{*B, C, G, I, J, K, O, R, S*}, and **UnknownNodeSet** is the set of remaining nodes which are unknown to the decision makers.

Step 2: Read X from **UnknownNodeSet**. If the value returned is **NULL**, then go to Step 6.

Step 3: Read the values of X 's neighbors. For example, when *E* is read, the values of *E*'s neighbors, *B* and *F*, are inspected. Similarly, when *N* is read, the values of *N*'s neighbors, *J*, $\widetilde{M}$, *O*, *Q*, are inspected.

Step 4: Compute $P(X=1 \mid w_X)/P(X=0 \mid w_X)$.

Step 5: Assign 0 or 1 to X from a random number generator favoring by the ratio $P(X=1 \mid w_X)/P(X=0 \mid w_X)$. Go to Step 2.

Step 6: Compute the belief of $X=1$, $BEL(x)$, from the proportion of 1's of X.

End of Algorithm 2.

For the propagation of the fuzzy nodes, we design the fuzzy simulation algorithm to generate the possibility belief range of the unknown fuzzy nodes. In Algorithm 3, we replace the operator "$\otimes$" with minimum intersection operator "$\wedge$".

<u>**Algorithm 3**</u>: **Fuzzy simulation for fuzzy nodes**

Step 1: Read **EvidenceSet**, **UnknownNodeSet**.

Step 2:Read $\widetilde{Y}$ from **UnknownNodeSet**. If the value returned is **NULL**, then go to Step 8.

Step 3: Read the values of $\widetilde{Y}$'s neighbors. For example, when $\widetilde{D}$ are read, the values of $\widetilde{D}$'s neighbors, *H*, are inspected.

Step 4: Generate $Pos(\widetilde{y} \mid w_{\widetilde{Y}}) = Pos(\widetilde{y}) \wedge Pos(\widetilde{\psi}) \wedge P(x \mid \pi_x)$, where $Pos(\widetilde{\psi})$ stands for the possibility distribution of $\widetilde{Y}$'s fuzzy neighbors, $P(x \mid \pi_x)$ stands for the probability distribution of $\widetilde{Y}$'s crisp neighbor. For $\widetilde{M}$, the crisp neighbors are *L* and *N*, and the fuzzy neighbors are empty.

Step 5: Specify the λ level and generate λ-cut set of $\widetilde{Y}$.

Step 6: Sample from the λ-cut set of $\widetilde{Y}$ randomly and get y.

Step 7: Store the minimum $MIN(y)$, the maximum $MAX(y)$, and the mean value $MEAN(y)$ of the sampled y. Go to Step 2.

Step 8: Generate the belief distribution of $\widetilde{Y}$, $BEL(\widetilde{y})$, in the form of ($MIN(y), MEAN(y), MAX(y)$).

End of Algorithm 3.

This study implements Algorithm 2 and 3 with Java 2. We observe the belief distribution for three periods and specify the confidence of the possibility functions at λ = 0.0, 0.5 and 1.0. The program simulates for 10000 iterations and finishes in less than 1 second. The results of simulation show the belief of every unknown node at every period, which are summarized in Table 14a through Table 14c. Observing Table 14a to Table 14c, we can find the belief distributions of the nodes relevant to poor schedule adherence. Several causes have significant influence on the supply chain inefficiency. The customer has large product variety ($A^t = 1$) with belief and around 0.8. The frequent specification alterations put by the customer

($E^t = 1$) is 0.55 to 0.65, approximately. Besides, the customer puts frequent schedule build alterations ($L^t = 1$) with belief 0.65, 0.88, 0.88, according to λ=0, 0.5 and 1.0, respectively. All the manifestations provide direct or indirect explanations for Company S's poor schedule adherence. The results also explain why the decision variable ICT aims at L and is expected to treat the supply chain through the effect of U on I. However, the single company is not capable of solving the problems in this two-echelon supply chain.

Next, Company S needs to initiate a collaborative plan with Company C to improve the schedule adherence and supply chain performance.

5.4.2 Phase II: Optimization phase

Now Company C realized what related to the poor schedule adherence in the two-echelon supply chain. It sensed the needs to implement electronic supply chain collaboration via ICT to upgrade the collaborative platform.

Table 14a: The results of simulation (λ -level = 0.0)

	t =1	t =2	t =3
Crisp nodes			
$BEL(A^t = 1)$	.8248	.8286	.8220
$BEL(E^t = 1)$	.6236	.6466	.6467
$BEL(F^t = 1)$	.3370	.3121	.3126
$BEL(H^t = 1)$	.5221	.5113	.5189
$BEL(L^t = 1)$	.6623	.6378	.6433
$BEL(N^t = 1)$	.4369	.4533	.4465
$BEL(Q^t = 1)$	.2845	.2965	.3003
Fuzzy nodes			
$BEL(\tilde{d}^t)$	(.3001, .5984, .8999)	(.3001, .6031, .8999)	(.3000, .6008, .8999)
$BEL(\tilde{m}^t)$	(.5001, .8185, .9999)	(.5000, .8132, .9999)	(.5000, .8164, .9999)
$BEL(\tilde{p}^t)$	(.5000, .6009, .6999)	(.5000, .5992, .6999)	(.5000, .6000, .6999)

Note: The evidence set $\dot{\mathbf{E}}$={ B^t =1, C =1, G^t =0, I^t =1, J^t =0, K^t =1, O^t =0, R^t =1, S^t =1| $1 \le t \le 3$ }.

Table 14b: The results of simulation (λ -level = 0.5)

	t =1	t =2	t =3
Crisp nodes			
$BEL(A^t=1)$	.8355	.8389	.8367
$BEL(E^t=1)$	.5587	.5323	.5599
$BEL(F^t=1)$	.4342	.4558	.4319
$BEL(H^t=1)$	.6395	.6541	.6530
$BEL(L^t=1)$	.8729	.8897	.8847
$BEL(N^t=1)$	.8121	.8429	.8442
$BEL(Q^t=1)$	.1816	.2775	.2783
Fuzzy nodes			
$BEL(\tilde{d}^t)$	(.4350, .5760, .7574)	(.4350, .5728, .757)	(.4350, .5738, .7574)
$BEL(\tilde{m}^t)$	(.5450, .6463, .9774)	(.5450, .6342, .9774)	(.5450, .6338, .9774)
$BEL(\tilde{p}^t)$	(.5250, .5935, .6549)	(.5250, .5897, .6549)	(.5250, .5899, .6549)

Note: The evidence set $\dot{\mathbf{E}}$={ B^t =1, C =1, G^t =0, I^t =1, J^t =0, K^t =1, O^t =0, R^t =1, S^t =1\| $1 \le t \le 3$ }.

Table 14c: The results of simulation (λ-level = 1.0):

	t =1	t =2	t =3
Crisp nodes			
$BEL(A^t=1)$	.8350	.8305	.8381
$BEL(E^t=1)$	.5410	.5389	.5458
$BEL(F^t=1)$	.4446	.4544	.4360
$BEL(H^t=1)$	.6291	.6760	.6470
$BEL(L^t=1)$	.8698	.8972	.8817
$BEL(N^t=1)$	.8419	.8620	.8520
$BEL(Q^t=1)$	.0779	.1064	.1046
Fuzzy nodes			
$BEL(\tilde{d}^t)$	(.5700, .5934, .6149)	(.5700, .5923, .6149)	(.5700, .5929, .6149)
$BEL(\tilde{m}^t)$	(.5900, .6518, .9549)	(.5900, .6445, .9549)	(.5900, .6481, .9549)
$BEL(\tilde{p}^t)$	(.5500, .6027, .6099)	(.5500, .6016, .6099)	(.5500, .6018, .6099)

Note: The evidence set $\dot{\mathbf{E}}$={B^t=1, C=1, G^t=0, I^t=1, J^t=0, K^t=1, O^t=0, R^t=1, S^t=1|$1 \le t \le 3$}.

After intensive communications, the two participating businesses finally reach a common consensus and start implementing this cooperative proposal. They start considering relevant Information technologies and collaborative solutions, especially business-to-business (B-to-B) collaboration expected to improve the information transparency and customer relationships in the supply chain value streams. After professional survey and consulting experts' opinions, four alternatives will be evaluated. The information for the optimization phase is listed as follow.

$ICT = \{z_1, z_2, z_3, z_4\}$, where z_1 stands for XML, z_2 stands for RosettaNet, z_3 stands for ebXML, and z_4 stands for BizTalk. All z_i are 0-1 variables. When the i th alternative is adopted, z_i is 1; otherwise z_i is 0. More than one alternative can be implemented simultaneously to cover various needs from different partners.

$U = \{U^1, U^2, U^3, U^4, U^5, U^6\}$: set of costs and utilities. Refer to Table 16.

$u_{i,j=1,2,3,4}$ Costs of z_i at phase j, where j=1 represents the conceptualization phase, j=2 represents the analysis phase, j=3 represents the design and implementation phase, j=4 represents the maintenance phase. Notably, the costs include all explicit expenditures and implicit efforts to implement the alternatives.

$u_{i,l=5,6}$: Expected benefits of z_i, where l=5 means information transparency via z_i, l=6 means expected on-time product delivery resulted from solution z_i.

All the measures of utility set are normalized to scale 0 to 100.

$\tilde{p}_v$: Fuzzy parameters that defines "around v", where $v \in \{40,50,60,70,80,90,95\}$.

$Pos(\tilde{p}_v)$: Possibility function of fuzzy parameter $\tilde{p}_v$, $\underline{p_v} \le p_v \le \overline{p}_v$, and $Pos(p_v^*)$=1.

Table 15: Description of IT solutions

	Description	Software provider
XML (z_1)	eXtensible Markup Language (XML) [53] is a simple and flexible text format derived from SGML. An important role in exchanging of data and basis of technologies, such as Web Services, SOAP and etc.	webMethods [54], Peregrine System [50], Ariba [45], TIBCO [42], etc.
RosettaNet (z_2)	RosettaNet is an e-business process standard for Information Technology, Electronic Components, Semiconductor Manufacturing and Telecommunications industries. Sponsored by IBM, Intel, Inovis, webMethods and over 400 companies. [51]	webMethods, Peregrine System, TIBCO, Ariba …
EbXML (z_3)	Electronic Business using eXtensible Markup Language (ebXML) [47] provides an open XML-based standard to support exchanging business messages and conduct trading relationships. Sponsored by UN/CEFACT and OASIS.	webMethods , Peregrine System, TIBCO, …
BizTalk (z_4)	BizTalk [46] is an XML-based and open standard sponsored by Microsoft Corporation. It is used to adopt the business processes of business-to business (B-to-B), enterprise application integration (EAI), and Business Process Automation (BPA).	Microsoft BizTalk Server [49]

The descriptions the four solutions are given in Table 15. Also, the costs to deploy every alternative, including all explicit expenditures and implicit efforts, are estimated in Table 16. Due to the uncertainty and ambiguity involved in the planning processes, all the costs and utilities of the ICT alternatives are expressed by fuzzy numbers. For convenience of comparison and to avoid the confidential issues about the price, the fuzzy numbers are normalized and standardized to [0,100]. The information

Table 16: Estimated costs/utilities of IT solutions in Example 4 (scale: 0 to 100)

Costs/Utilities	XML (z_1)	RosettaNet(z_2)	EbXML(z_3)	BizTalk(z_4)
Cost of Conceptualization ($u_{i,1}$)	$\tilde{p}_{40}$	$\tilde{p}_{90}$	$\tilde{p}_{70}$	$\tilde{p}_{60}$
Cost of Analysis ($u_{i,2}$)	$\tilde{p}_{50}$	$\tilde{p}_{90}$	$\tilde{p}_{80}$	$\tilde{p}_{70}$
Cost of Design ($u_{i,3}$)	$\tilde{p}_{50}$	$\tilde{p}_{90}$	$\tilde{p}_{70}$	$\tilde{p}_{60}$
Cost of Maintenance ($u_{i,4}$)	$\tilde{p}_{50}$	$\tilde{p}_{80}$	$\tilde{p}_{60}$	$\tilde{p}_{60}$
Information transparency (u_i)	$\tilde{p}_{60}$	$\tilde{p}_{95}$	$\tilde{p}_{60}$	$\tilde{p}_{80}$
Reduced forecast variance (v_i)	$\tilde{p}_{50}$	$\tilde{p}_{90}$	$\tilde{p}_{60}$	$\tilde{p}_{80}$

By proposition 1, we can express all the fuzzy parameters in costs and utilities as in Table 17. Furthermore, by Proposition 2, the non-linear membership functions are transformed into equivalent linear functions.

Table 17: The possibility functions of the fuzzy parameters in Example 4

$\tilde{p}_k$	Possibility function $Pos(\tilde{p}_k)$	$(\underline{p}_v, p_v^*, \overline{p}_v)$
$\tilde{p}_{40}$	$0.2(p_{40}-35)-0.2(\lvert p_{40}-40\rvert+p_{40}-40)$	(35,40,45)
$\tilde{p}_{50}$	$0.2(p_{50}-45)-0.2(\lvert p_{50}-50\rvert+p_{50}-50)$	(45,50,55)
$\tilde{p}_{60}$	$0.2(p_{60}-55)-0.2(\lvert p_{60}-60\rvert+p_{60}-60)$	(55,60,65)
$\tilde{p}_{70}$	$0.2(p_{70}-65)-0.2(\lvert p_{70}-70\rvert+p_{70}-70)$	(65,70,75)
$\tilde{p}_{80}$	$0.2(p_{80}-75)-0.2(\lvert p_{80}-80\rvert+p_{80}-80)$	(75,80,85)
$\tilde{p}_{90}$	$0.2(p_{90}-85)-0.2(\lvert p_{90}-90\rvert+p_{90}-90)$	(85,90,95)
$\tilde{p}_{95}$	$0.2(p_{95}-90)-0.2(\lvert p_{95}-95\rvert+p_{95}-95)$	(90,95,100)

In the optimization phase, there are four objectives to be optimized: maximizing the expected information transparency (obj_1), maximizing the expected customer satisfaction (obj_2), minimizing the expected costs (obj_3), and maximizing the overall membership of the fuzzy parameters (obj_4). All the four alternatives are possibly selected to implement. The lower bound of expected information transparency and reduced forecast variance are set at 100. Oppositely, the upper limit of costs is 650 units. Now the model for supply chain collaboration is formulated as Model 2.

Model 2

$$
\begin{aligned}
&Max \quad obj_1 = \sum_{i=1}^{4} z_i u_{i,5} = z_1\tilde{p}_{60} + z_2\tilde{p}_{95} + z_3\tilde{p}_{60} + z_4\tilde{p}_{80}, \\
&Max \quad obj_2 = \sum_{i=1}^{4} z_i u_{i,6} = z_1\tilde{p}_{50} + z_2\tilde{p}_{90} + z_3\tilde{p}_{60} + z_4\tilde{p}_{80}, \\
&Min \quad obj_3 = \sum_{i=1}^{4}\sum_{j=1}^{4} z_i u_{i,j} = (z_1\tilde{p}_{40} + 3z_1\tilde{p}_{50}) + (3z_2\tilde{p}_{90} + z_2\tilde{p}_{80}) \\
&\qquad\qquad + (2z_3\tilde{p}_{70} + z_3\tilde{p}_{60} + z_3\tilde{p}_{80}) + (3z_4\tilde{p}_{60} + z_4\tilde{p}_{70}), \\
&Max \quad obj_4 = Pos(\tilde{p}_v)
\end{aligned}
\tag{48}
$$

s.t.

$$
\begin{aligned}
&Pos(\tilde{p}_{40}) = 0.2(p_{40} - 35) - 0.2(|p_{40} - 40| + p_{40} - 40), \\
&Pos(\tilde{p}_{50}) = 0.2(p_{50} - 45) - 0.2(|p_{50} - 50| + p_{50} - 50), \\
&Pos(\tilde{p}_{60}) = 0.2(p_{60} - 55) - 0.2(|p_{60} - 60| + p_{60} - 60), \\
&Pos(\tilde{p}_{70}) = 0.2(p_{70} - 65) - 0.2(|p_{70} - 70| + p_{70} - 70), \\
&Pos(\tilde{p}_{80}) = 0.2(p_{80} - 75) - 0.2(|p_{80} - 80| + p_{80} - 80), \\
&Pos(\tilde{p}_{90}) = 0.2(p_{90} - 85) - 0.2(|p_{90} - 90| + p_{90} - 90), \\
&Pos(\tilde{p}_{95}) = 0.2(p_{95} - 90) - 0.2(|p_{95} - 95| + p_{95} - 95), \\
&\underline{p}_v \le p_v \le \overline{p}_v, \\
&z_i = 0 \quad or \quad 1, \\
&obj_1 \ge 100, \quad obj_2 \ge 100, \quad obj_3 \le 650, \\
&Pos(\tilde{p}_v) \ge 0.5.
\end{aligned}
\tag{49}
$$

This study specifies the confidence level of fuzzy sets at 0.5 as in the last equalities in (49). It regulates that every possibility must be equal to or greater than 0.5, which excludes the case when the costs and utilities are poorly estimated. Since the four objectives are nonlinear functions, the global optimal solutions will not be solved directly. So, this study converts these nonlinear functions into linear ones with the linearization strategies in next subsection.

5.4.3 Linearization strategies

The following constraints can convert the nonlinear non-separate term $z_i p_j$ in the first three objectives into linear ones, where z_i is a 0-1 variable and p_j is a continuous variable. First, replace $z_i p_j$ with $q_{i,j}$. The behaviors of $q_{i,j}$ can be bounded with a set of linear constraints as follow.

$$\begin{aligned} & p_j + \overline{p}_j (z_i - 1) \le q_{i,j} \le p_j + \overline{p}_j (1 - z_i), \\ & 0 \le q_{i,j} \le z_i \overline{p}_j. \end{aligned} \tag{50}$$

The authors then verify the manifestation of $q_{i,j}$ with the instances of z_i. When $z_i = 0$, (50) will be changed into (51) which implies that $q_{i,j} = 0$.

$$\begin{aligned} & p_j - \overline{p}_j \le q_{i,j} \le p_j + \overline{p}_j, \\ & 0 \le q_{i,j} \le 0. \end{aligned} \tag{51}$$

In the other case when $z_i = 1$, (50) will change into (52) which implies that $q_{i,j} = z_i \overline{p}_j$.

$$\begin{aligned} & p_j \le q_{i,j} \le p_j, \\ & 0 \le q_{i,j} \le z_i \overline{p}_j. \end{aligned} \tag{52}$$

Linearizing the first three objectives with (50) and the fourth objective with **Proposition 2**, **Model 2** can be converted into **Model 3**.

Model 3

$$
\begin{aligned}
&Max \quad obj_1 = q_{1,60} + q_{2,95} + q_{3,60} + q_{4,80}, \\
&Max \quad obj_2 = q_{1,50} + q_{2,90} + q_{3,60} + q_{4,80}, \\
&Min \quad obj_3 = (q_{1,40} + 3q_{1,50}) + (3q_{2,90} + q_{2,80}) + (2q_{3,70} + q_{3,60} + q_{3,80}) \\
&\qquad\qquad + (3q_{4,60} + q_{4,70}), \\
&Max \quad obj_4 = \sum_i Pos(\tilde{p}_i) = Pos(\tilde{p}_{40}) + Pos(\tilde{p}_{50}) + Pos(\tilde{p}_{60}) + Pos(\tilde{p}_{70}) \\
&\qquad\qquad + Pos(\tilde{p}_{80}) + Pos(\tilde{p}_{90}) + Pos(\tilde{p}_{95})
\end{aligned} \tag{53}
$$

$$
\begin{aligned}
&p_{40} + \bar{p}_{40}(z_1 - 1) \le q_{1,40} \le p_{40} + \bar{p}_{40}(1 - z_1), \qquad 0 \le q_{1,40} \le z_1 \bar{p}_{40}, \\
&p_{50} + \bar{p}_{50}(z_1 - 1) \le q_{1,50} \le p_{50} + \bar{p}_{50}(1 - z_1), \qquad 0 \le q_{1,50} \le z_1 \bar{p}_{50}, \\
&p_{60} + \bar{p}_{60}(z_1 - 1) \le q_{1,60} \le p_{60} + \bar{p}_{60}(1 - z_1), \qquad 0 \le q_{1,60} \le z_1 \bar{p}_{60}, \\
&p_{80} + \bar{p}_{80}(z_2 - 1) \le q_{2,80} \le p_{80} + \bar{p}_{80}(1 - z_2), \qquad 0 \le q_{2,80} \le z_2 \bar{p}_{80}, \\
&p_{90} + \bar{p}_{90}(z_2 - 1) \le q_{2,90} \le p_{90} + \bar{p}_{90}(1 - z_2), \qquad 0 \le q_{2,90} \le z_2 \bar{p}_{90}, \\
&p_{95} + \bar{p}_{95}(z_2 - 1) \le q_{2,95} \le p_{95} + \bar{p}_{95}(1 - z_2), \qquad 0 \le q_{2,95} \le z_2 \bar{p}_{95}, \\
&p_{60} + \bar{p}_{60}(z_3 - 1) \le q_{3,60} \le p_{60} + \bar{p}_{60}(1 - z_3), \qquad 0 \le q_{3,60} \le z_3 \bar{p}_{60}, \\
&p_{70} + \bar{p}_{70}(z_3 - 1) \le q_{3,70} \le p_{70} + \bar{p}_{70}(1 - z_3), \qquad 0 \le q_{3,70} \le z_3 \bar{p}_{70}, \\
&p_{80} + \bar{p}_{80}(z_3 - 1) \le q_{3,80} \le p_{80} + \bar{p}_{80}(1 - z_3), \qquad 0 \le q_{3,80} \le z_3 \bar{p}_{80}, \\
&p_{60} + \bar{p}_{60}(z_4 - 1) \le q_{4,60} \le p_{60} + \bar{p}_{60}(1 - z_4), \qquad 0 \le q_{4,60} \le z_4 \bar{p}_{60}, \\
&p_{70} + \bar{p}_{70}(z_4 - 1) \le q_{4,70} \le p_{70} + \bar{p}_{70}(1 - z_4), \qquad 0 \le q_{4,70} \le z_4 \bar{p}_{70}, \\
&p_{80} + \bar{p}_{80}(z_4 - 1) \le q_{4,80} \le p_{80} + \bar{p}_{80}(1 - z_4), \qquad 0 \le q_{4,80} \le z_4 \bar{p}_{80},
\end{aligned} \tag{54}
$$

$$
\begin{aligned}
&Pos(\tilde{p}_{40}) = 0.2(p_{40} - 35) - 2[0.2(p_{40} - 40 + d_{40})], \\
&Pos(\tilde{p}_{50}) = 0.2(p_{50} - 45) - 2[0.2(p_{50} - 50 + d_{50})], \\
&Pos(\tilde{p}_{60}) = 0.2(p_{60} - 55) - 2[0.2(p_{60} - 60 + d_{60})], \\
&Pos(\tilde{p}_{70}) = 0.2(p_{70} - 65) - 2[0.2(p_{70} - 70 + d_{70})], \\
&Pos(\tilde{p}_{80}) = 0.2(p_{80} - 75) - 2[0.2(p_{80} - 80 + d_{80})], \\
&Pos(\tilde{p}_{90}) = 0.2(p_{90} - 85) - 2[0.2(p_{90} - 90 + d_{90})], \\
&Pos(\tilde{p}_{95}) = 0.2(p_{95} - 90) - 2[0.2(p_{95} - 95 + d_{95})], \\
&\underline{p}_v \le p_v \le \bar{p}_v, \\
&z_i = 1 \quad or \quad 0, \\
&obj_1 \ge 100, \quad obj_2 \ge 100, \quad obj_3 \le 650, \\
&p_v + d_v \ge p_v^*, \quad 0 \le d_v \le p_v^*, \\
&Pos(\tilde{p}_v) \ge 0.5.
\end{aligned} \tag{55}
$$

This study uses modified Zimmermann's approach [22,44] to solve this fuzzy multi-objective decision model. The steps for solving **Model 3** are as follow.

Step 1: Solve all objectives for their ideal and anti-ideal solutions separately and independently. We maximize and minimize obj_1 to get its ideal and anti-ideal solution denoted by obj_1^* and obj_1^-, respectively; obj_2^*, obj_2^-, obj_3^*, obj_3^-, obj_4^* and obj_4^- are obtained in the same way. Notably, obj_3 is minimized for its ideal solution and maximized for its anti-ideal solution.

Step 2: Maximize the average score subjected to each score of individual objective and the original constraints of the optimization program. The program will be converted into (56).

$$\begin{aligned}
&Max \quad score \\
&s.t. \\
&\quad score \le (score_1 + score_2 + score_3 + score_4)/4, \\
&\quad score_1 = \frac{obj_1 - obj_1^-}{obj_1^* - obj_1^-}, \quad score_2 = \frac{obj_2 - obj_2^-}{obj_2^* - obj_2^-}, \\
&\quad score_3 = \frac{obj_3 - obj_3^-}{obj_3^* - obj_3^-}, \quad score_4 = \frac{obj_4 - obj_4^-}{obj_4^* - obj_4^-}, \\
&\quad (54) - (55).
\end{aligned} \tag{56}$$

In **Example 4**, the numerical instance for (56) is (57).

$$
\begin{aligned}
& Max \quad score \\
& s.t. \\
& \quad score \le \left(score_1 + score_2 + score_3 + score_4\right)/4, \\
& \quad score_1 = \frac{obj_1 - 110}{185 - 110}, \qquad score_2 = \frac{obj_2 - 100}{180 - 100}, \\
& \quad score_3 = \frac{obj_3 - 650}{400 - 650}, \qquad score_4 = \frac{obj_4 - 0}{7 - 0}, \\
& \quad (54) - (55).
\end{aligned} \tag{57}
$$

LINGO 8 solves the above program and finds the global optimal solutions in less than 0.1 second.

From the solution report, the optimal solution for supply chain collaboration is to implement RosettaNet (z_2) and BizTalk (z_4), which yield a overall score of 0.7354, the expected information transparency of 175 standard units, the expected on-time product delivery of 170 standard units, the total costs of 600 units, and the overall membership of 7.0 (for seven fuzzy parameters). The individual scores of the information transparency, customer satisfaction, total costs and possibilities of fuzzy parameters are 0.867, 0.875, 0.200, and 1.000, respectively. The detailed solution report is listed in Table 18.

5.5 Discussions and conclusions

This paper proposes an integrated model for supply chain diagnostics and treatment optimization. The authors adopt dynamic fuzzy influence diagrams to describe the cause-and-effect relationships in the two-echelon supply chain. In addition to the random nodes standing for the key variables in the business practice, we use one decision node representing the treatment to the supply chain problems and a value node standing for the objectives to be optimized. In the dynamic fuzzy influence diagrams, two kinds of nodes are designed: crisp discrete nodes and fuzzy continuous nodes. In living expert systems, the knowledge engineers can maintain and update the distributions of the nodes via a system interface as needed. The decision-makers can conduct diagnostic reasoning based on the observed symptoms or evidences for multiple periods. Then the optimal solution to treat the diagnosed problem will be suggested by the fuzzy multi-objective optimization models.

Table 18: Solution report of Example 4

Objective value = 0.7354		
$score = 0.7354$	$p_{40} = 40.0000$	$Pos(\tilde{p}_{40}) = 1.0000$
$score_1 = 0.8667$	$p_{50} = 50.0000$	$Pos(\tilde{p}_{50}) = 1.0000$
$score_2 = 0.8750$	$p_{60} = 60.0000$	$Pos(\tilde{p}_{60}) = 1.0000$
$score_3 = 0.2000$	$p_{70} = 70.0000$	$Pos(\tilde{p}_{70}) = 1.0000$
$score_4 = 1.0000$	$p_{80} = 80.0000$	$Pos(\tilde{p}_{80}) = 1.0000$
$obj_1 = 175.0000$	$p_{90} = 90.0000$	$Pos(\tilde{p}_{90}) = 1.0000$
$obj_2 = 170.0000$	$p_{95} = 95.0000$	$Pos(\tilde{p}_{95}) = 1.0000$
$obj_3 = 600.6667$		
$obj_4 = 7.0000$		
$z_2 = z_4 = 1$		
$z_1 = z_3 = 0$		

In implementing the influence diagram, we encounter a dilemma. The influence diagrams are compact and descriptive in model decision model conceptually. However, when we operate the model, there is time delay between diagnosis and treatment. The conventional methods in evaluating influence diagram did not consider the time delay. They even did not consider the diagnosis but concentrate on maximizing the expected gain. Hence, this study divides the reasoning procedure into two phases: diagnostic and optimization phases. The decision-makers are allowed to check the diagnosed problems first and then decide what the optimal solution to the problem is. This division makes the operations of the influence diagram more consistent in real-world industrial practices.

This study is a proposed design of integrated framework for supply chain diagnosis and decision-making. For future extensions, this work may be implemented into complete decision support systems. In the decision support systems, various models and reasoning strategies can be included in the model bases. Through the user interface, the decision-makers may modify the influence diagram structures and related parameters. The probability distributions and possibility distributions can be learned and tuned in light of different scenarios or constraints.

Chapter 6 Discussions and conclusions

This dissertation defines general Bayesian networks (GBN) that are composed of five components: the set of crisp nodes, the set of fuzzy nodes, the set of crisp parameters, the set of fuzzy parameters, the set of arcs (links) among the nodes, and the set of decision variables. Three categories of reasoning are studied as the special cases (subsets) of general Bayesian networks: (1) diagnosis with crisp nodes and fuzzy parameters, (2) diagnosis and decision-making with crisp nodes and fuzzy parameters, and (3) diagnosis and decision-making with fuzzy nodes in dynamic environments. The distinguished features of this dissertation include:

1. Define general Bayesian networks as the general research framework.
2. Solves the reasoning tasks in three subsets of GBN where different types/degrees of uncertainties are considered.
3. Consider extra knowledge or constraints for the belief propagation, which are not implemented in the formal knowledge bases.
4. Answer the queries from Bayesian networks in dynamic as well as static environments.

This chapter will first discuss the implications from the series of research, give some directions for future extensions, and make the conclusions.

6.1 Discussions

In implementing the dissertation, the author finds some issues worth further discussion.

First, when the decision variables are introduced into the Bayesian networks, the models will become influence diagrams. However, when the reasoning tasks include diagnosis and follow-up treatments, the author encounters the dilemma in using the decision model. The dilemma results from the time delay between diagnosis and decision-making. When an expert diagnoses a problem, supposed he or she has not ascertained where the symptoms originate, not to mention the assessment of different solution. The roots of problems determine the set of alternative solutions and their outcomes. That is what the author calls time delay between diagnosis and decision-making. However, in using influence diagrams, the decision-makers have to estimate the costs and utilities of every alternative, regardless how the roots influence determine the solution sets and their values. That is why traditional approaches evaluating influence diagrams focused only on maximizing the expected gain (utility) and ignoring the diagnostic reasoning. Hence, in Problem 4 (Chapter 4), this dissertation partitions the reasoning in influence diagrams into two phases: diagnostic phase and decision-making phase.

Second, in handling the fuzzy parameter or fuzzy variables, this dissertation uses two approaches: piecewise linearization and α-cut methods. The advantage of piecewise linearization is the quality and performance in solving the reasoning model, especially when the reasoning model is designed as a nonlinear programming model. However, when the problem scale grows large, the programming model may be too complex. Under such a circumstance, α-cut methods perform better, especially for fuzzy simulation.

Third, in this dissertation, the conditional distributions of the nodes with fuzzy parents are simplified by partition methods, in which the distributions of the nodes are conditioned on some sub-domains of their fuzzy parents. There is another alternative approach for the conditional distributions: functional distributions. That is, the distributions of the nodes can be expressed as a function of their fuzzy parents' values. The functional distributions need more complex computation schema but are more compact and logical consistent.

6.2 Future extensions

For extended studies, there are several interesting issues worth further investigation.

First, the proposed models can be extended to complete decision support systems (DSS). A DSS is composed of four subsystems: data management subsystem, mode management subsystem, knowledge-based subsystem, and user interfaces. The schematic view of DSS is show in Figure 12. To apply the models into real-world application, the real databases can be integrated with the reasoning systems, and assess the validity and reliability of the proposed models.

Second, for enhancing the computational efficiency, some algorithmic methods can be combined into the models, such as artificial neural networks, genetic algorithms, and so on. The outcomes from different computation schemes can be compared and cross-referenced.

Third, several decision-making methods can be integrated with the proposed models, especial the multiple criteria decision making (MCDM) methods, e.g. Analytic hierarchy process (AHP), Data encryption analysis (DEA), and so on. These decision-making methods can provide optimal treatment or approaches to the diagnosed problems and enhance the decision quality in the reasoning models.

This dissertation intends to contribute to diagnostic reasoning in both methodology and applications, especially in industrial practices and medical informatics.

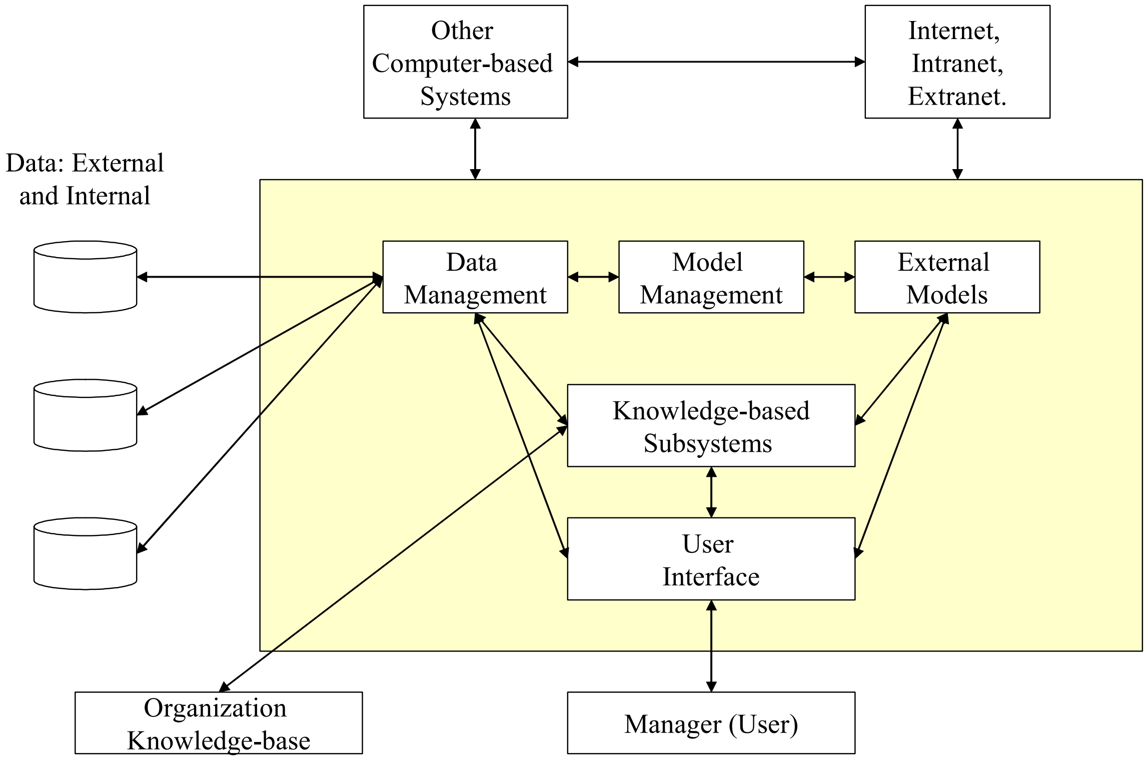

Figure 12: A schematic view of DSS

6.3 Concluding remarks

This dissertation proposes a definition of general Bayesian networks, which can be specialized into various kinds of Bayesian networks. The general Bayesian networks provide a foundation stone for flexible and robust knowledge base design. The knowledge base can solve various problems involving fuzzy as well as crisp information, under dynamic as well as static circumstances.

This dissertation solves three categories of reasoning are studied as the special cases (subsets) of general Bayesian networks: (1) diagnosis with crisp nodes and fuzzy parameters, (2) diagnosis and decision-making with crisp nodes and fuzzy parameters, and (3) diagnosis and decision-making with fuzzy nodes in dynamic environments. If taking the costs and utilities in Problem 4 (Chapter 4) as one type of fuzzy parameters, this dissertation has develop a whole model for solving the general Bayesian networks. The author hopes that this dissertation has make some contributions in expert systems and reasoning methods.

References

[1] Castill E., Gutoerre J. M., Hadi A. S., "A new method for symbolic inference in Bayesian networks", Networks, Vol. 28, 1996, 31-43.

[2] Castillo E., Gutoerrez J. M., Hadi A. S., Expert systems and probabilistic network models, Spinger-Verlag New York, Inc., 1997.

[3] Choi, T. Y., Z. Wu, L. Ellram, and B. R. Koka, Supplier-supplier relationships and their implications for buyer-supplier relationships, IEEE transactions on engineering management, 49 (2) (2002) 119-130.

[4] Dagum, P., A. Galper & E. Horvitz, Dynamic network models for forecasting. Proceedings of the 8th conference on uncertainty in Artificial intelligence (UAI' 92), Stanford University, San Francisco (CA), Morgan Kaufmann (1992) 41-48.

[5] Ecker, J. G., M. Kupferschmid, C. E. Lawrence, A. A. Reilly, and A. C. H. Scott, An application of nonlinear optimization in molecular biology, European journal of operational research, 138 (2002) 452-458.

[6] Floudas C. A., "Deterministic global optimization- theory and applications," Kluwer Academic Publishers, 2000.

[7] Freeling, A. N. S., Possibility versus fuzzy probabilities- two alternative decision aids, Fuzzy sets and decision analysis, TIMS/ Studies in the management sciences, 20 (1984) 67-81.

[8] Galán, S. F., F. Aguado, F. J. Díez & J. Mira, NasoNet, modeling the spread of nasopharyngeal cencer with networks of probabilistic events in discrete time. Artificial Intelligence in Medicine, 25 (2002) 247-264.

[9] Giannoccaro, I., P. Pontrandolfo and B. Scozzi, A fuzzy echelon approach for inventory management in supply chains, European Journal of Operational Research, 149 (2003) 185-196.

[10] Heide, J. B. and G. John, Alliances in industrial purchasing: The determinants of joint action in buyer-supplier relationships, Journal of Marketing Research, 27 (1) (1990) 24-36.

[11] Howard R. A. and Matheson, J. E. (1981). Influence diagrams. In Principles and applications of decision analysis, vol. 2, (1984). Menlo Park, Calif.: Strategic Decisions Group.

[12] Holm, D. B., K. Eriksson, and J. Johanson, "Creating value through mutual commitment to business network relationships," Strategic Management Journal, 20 (5) (1999) 467-486.

[13] Huang, C-H, H-Y Kao & H-L Li, "Web Services for forward integration in international tourism supply chains: a case study of tourism in Thailand", International Conference on Electronic Business, Singapore, Dec. 9-13, 2003.

[14] Kao, H-Y and H-L Li, "A diagnostic reasoning and optimal treatment model for Bacterial infections with fuzzy information," Computer Methods and Programs in Biomedicine, (revised and under review) 2003.

[15] Kao, H-Y, C-H Huang & H-L Li, "An Optimization Model for Dynamic Supply Chain Diagnostics and Decision-Making", European Applied Business Research Conference, Venice, Italy, Jun. 9-13, 2003.

[16] Kao, H-Y, C-H Huang and H-L Li, "Supply chain diagnostics with Dynamic Bayesian networks," accepted to Computers and Industrial Engineering, 2003.

[17] Kao, H-Y, H-L Li, and C-H Huang, "An integrated diagnostic and decision model of supply chain management", submitted, 2004.

[18] Kao, H-Y, J-J Shuai and H-L Li, "Probabilistic Reasoning in Supply Chain Management," The 5th International Conference on Industrial Engineering-Theory, Application and Practice, Taiwan, Dec. 10-12, 2000.

[19] Kaynak, H., "The relationship between just-in-time purchasing techniques and firm

performance”, IEEE transactions on engineering management, 49 (3) (2002) 205-217.

[20] Kononenko I., "Machine learning for medical diagnosis: history, state of the art and perspective", Artificial Intelligence in Medicine, 23, 2001, 89-109.

[21] Kurokawa, S., Make-or-buy decisions in R&D: small technology based firms in the United States and Japan, IEEE transactions on engineering management, (1997) Vol. 44, No. 2, 124-134.

[22] Lee, E. S. and R. J. Li, Fuzzy multiple objective programming and compromise programming with Pareto optimum, Fuzzy sets and systems, 53 (1993) 275-288.

[23] Leibovici Leonard, Fishman M., Schonheyder H. C., C. Riekehr, B. Kristensen, I. Shraga, Andreassen S., "A causal probabilistic network for optimal treatment of bacterial infections", IEEE Transactions on Knowledge and Data Engineering, Vol. 12, No. 4, 2000, 517-528.

[24] Li H-L, Chang C-T, Tsai J-F, “Approximately global optimization for assortment problems using piecewise linearization techniques”, European Journal of Operational Research, 140, 2002, 584-589.

[25] Li, H-L and H-Y Kao, "Causal reasoning from imperfect experiments and fuzzy information," accepted to International Journal of Fuzzy Systems, Vol. 6, No. 1, 2004, 44-51.

[26] Li, H-L and H-Y Kao, "Constrained Abductive Reasoning with Fuzzy Parameters in Bayesian networks," accepted to Computers & Operations Research, 2003

[27] Long W. J., "Medical informatics: reasoning methods", Artificial Intelligence in Medicine, 23, 2001, 71-87.

[28] McCullen, P. and D. Towill, Diagnosis and reduction of bullwhip in supply chains. Supply chain management, 7(3/4) (2002) 164-179.

[29] Naim, M. M., P. Childerhouse, S. M. Disney& D. R. Towill, A supply chain diagnostic methodology: determining the vector of change, Computers and Industrial Engineering,

43 (2002) 135-157.

[30] Nease, R. F. and D. K. Owens, Use of influence diagrams to structure medical decisions, medical decision making, 17 (1997) 263-275.

[31] Oliver R.M., Smith J.Q., Influence diagrams, belief nets and decision analysis, John Wiley & Sons, 1990.

[32] Olsen, R. F. and L. M. Ellram, "A portfolio approach to supplier relationships," Industrial Marketing Management, 26 (2) (1997) 101-113,.

[33] Owens, D.K., R.D. Shacher, and R. F. Nease, Representation and analysis of medical decision problems with influence diagrams, medical decision making, 17 (1997) 251-262.

[34] Pearl J., Causality-models, reasoning, and inference, Cambridge University Press, 2000.

[35] Pearl, J., Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan Kaufmann Publishers, Inc., 1997.

[36] Pool D., "Average-case analysis of a search algorithm for estimating prior and posterior probabilities in Bayesian networks with extreme probabilities," Proceedings of the 13th international joint conference on artificial intelligence, 13, 1, 1993, 606-612.

[37] Polya, G. 1954 Mathematics and plausible reasoning. Princeton: Princeton University Press.

[38] Shacher R. D., D'Ambrosi B., DelFabero B., "Symbolic probabilistic inference in belief networks," in Proceedings of the 10th conference on uncertainty in artificial intelligence, Morgan Kaufmann Publishers, San Francisco, 514-522.

[39] Tatman, J.A. and R. D. Shacher, Dynamic programming and influence diagrams, IEEE Transactions on systems, man, and cybernetics, 20(2) (1990) 365-379.

[40] Tsai, J-F, H-L Li and N-Z Hu, Global optimization for signomial discrete programming problems in engineering design, Engineering Optimization, 34(6) 2002.

[41] Winston, W. L. (1999), "Operations Research: applications and algorithms", 3rd,

Duxbury Press.

[42] Youssef, M. A., Agile manufacturing: a necessary condition competing in global markets, Industrial Engineering, 24 (12) (1992) 18-20.

[43] Zadeh, L. A., Fuzzy sets, Information and Control, 8 (1965), 338-353.

[44] Zimmermann, H-J, Fuzzy programming and linear programming with several objective functions, Fuzzy sets and systems, 1 (1978) 45-55.

[45] Ariba: http://www.ariba.com/

[46] BizTalk: http://www.biztalk.org/

[47] ebXML: http://www.ebxml.org/

[48] LINGO, http://www.lindo.com/cgi/frameset.cgi?leftlingo.html;lingof.html.

[49] Microsoft BizTalk Server: http://www.microsoft.com/biztalk/

[50] Peregrine System: http://www.inovis.com/

[51] RosettaNet Consortium: http://www.rosettanet.org/

[52] TIBCO: http://www.tibco.com/

[53] W3C, World Wide Web Consortium: http://www.w3.org/

[54] webMethods: http://www.webmethods.com/