

# Cyclic Bayesian Network for Software Project Iterative Process

Guoping Jiang<sup>1</sup> and Mingkui Ren<sup>2,\*</sup>

<sup>1,2</sup>Department of Management Engineering and Equipment Economy, Naval Engineering University, Wuhan Hubei, China, 430033

\*Corresponding author

**Abstract**— Aim at the Cyclic Bayesian Network, discusses the convergency of the ring's probability distribution. Simplify a cycle to a Simple-Cycle via node elimination operation, probability convergency of the cycle is proved. Based on cycle's probability convergency, probability propagation method is proposed.

**Keywords**—*bayesian network; cycle; convergency; probability propagation*

## I. INTRODUCTION

Software engineering practice shows<sup>[1]</sup>, software project's risks root in the uncertainties of software project process. Modeling software project process to abstract, represent and analyze the software process, can identify risk factors, estimate risk expense and construct risk relaxation program.

A Bayesian Network<sup>[2]</sup> (BN) is a graphical model that encodes relationship among variables of interest. At the same time, a BN is also a probability inference technique to deal with conditional uncertainty among different knowledge components, and it offers an ideal mathematic formalization for representation of uncertainty. A BN is the integration of graph theory and probability theory. A BN is a complicated evaluated causality graph intuitively, each node denotes a variable in the domain, namely event. The arc between nodes denotes the direct causality between events. Generally speaking, a BN is a directed acyclic graph  $(V, E)$ , where  $V$  is the set of nodes and  $E$  is the set of arcs. For any  $(x, y) \in E$ ,  $x$  is intitled as  $y$ 's parent node, marked as  $pa(y)$ , and  $y$  is  $x$ 's descent node, marked as  $desc(x)$ . Every node has its conditional probability distribution reflecting the dependence relationship between it's parent nodes and itself. A BN represents joint probability distribution of involved variables in the interested domain essentially. Due to the conditional independence, joint probability distribution of the network can be obtained by the chain equation (1). Lots of probability inference methods<sup>[3]</sup> have been exploited based on equation (1).

$$P = \pi_{i=1}^n p(x_i | pa(x_i)) \quad (1)$$

For the firm mathematic theoretic base, BN is deemed as the ideal tool for knowledge representation, inference and estimation in uncertain environment<sup>[4]</sup>. In the software engineering domain, BN has been used to estimate software's quality and faults successfully<sup>[5,6]</sup>. But when we model

software project process as a BN to estimate and evaluate software project risk, there are always directed cycles in the BN, for doing poorly done work over again, iterative development and etc. are possible in software project process. Probability inference of BN with directed cycles can not be carried out as Equation (1) shows. In this paper, we will discuss the problem of cyclic BN.

An outline of the remainder of the paper is as follows. In Section 2, we discuss how the cyclic BN can be simplified. In Section 3, we discuss the probability convergency for directed cycle. In section4, we discuss probability propagation for cyclic BN and carry out a case study. Finally, in Section 5, we conclude with a summary and a statement of future research.

## II. SIMPLIFY THE CYCLIC BAYESIAN NETWORKS

### A. Definition

Firstly, we suppose that all the nodes in this paper are discrete variables, and the number of their values is limited. The relative conceptions are listed as follows:

**Cyclic Bayesian Network**<sup>[7]</sup> (Cyclic BN): if any directed cycles existed in a Bayesian Network  $BN(V, E)$ , then  $BN(V, E)$  is named as a Cyclic Bayesian Network. Directed cycle is marked as  $L(N, R)$ , where  $N \subset E$  is the set of nodes on the cycle, and  $R \subset E$  is the set of directed arcs.

**Input of a cycle:** A node on a cycle  $v \in N$  which have parent nodes  $pa(v)$ , and  $pa(v)$  don't belong to the cycle  $pa(v) \notin N$ , then  $pa(v)$  are named as input nodes of this cycle and  $v$  as an Inceptor-Node.

**Inquired-Node:** An Inquired-Node is the node whose values and probability distribution demand to be obtained. For a directed cycle, the node possessing descent nodes which don't belong to this cycle, can be treated as an Inquired-Node.

**Node-Elimination**<sup>[7]</sup>: the operation of eliminating some node on a cyclic BN, at the same time changing the probability distribution of relative nodes conforming to certain rules, and having no influence on the probability propagation, is named as Node-Elimination.

**Normal-Node:** a node on a directed cycle is named as an Ordinary-Node, when it isn't an evidence node nor an Inquired-Node.

**Simple-Cycle:** a cycle possessing just two nodes is named as a Simple-Cycle.

*B. Simplify the Directed Cycle*

**Lemma 1**<sup>[7]</sup> For a cycle possessing no inputs, the Ordinary-Nodes can be eliminated one by one until the cycle is converted into a Simple-Cycle.

**Lemma 2** For a cycle possessing inputs, the Normal-Nodes can be eliminated one by one until the cycle is converted into a Simple-Cycle.

Proof: Let us consider a cycle possessing three nodes. There are three cases according to different Inquired-Node:

1) if the Inquired-Node is an evidence node

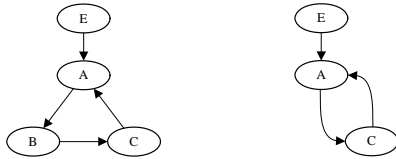


FIGURE I. IF INCEPTOR-NODE IS EVIDENCE NODE

As Figure 1 shows, Inceptor-Node *A* is an evidence node. After node *B* is eliminated, the topological structure is the right figure in Figure 1. Calculate each node's probability distribution as follows:

$$\begin{aligned} P(C) &= \sum_A \sum_E P'(A, E, C) \\ &= \sum_A \sum_B \sum_E P(A, B, C, E) \\ &= \sum_A \sum_E P(A | E) P(E) \sum_B P(B | A) P(C | B) \end{aligned}$$

Then  $P'(C | A) = \sum_B P(B | A) P(C | B)$ ,  $P'(A | C) = P(A | C)$  and

$$P'(A | E) = P(A | E).$$

2) If the Inceptor-Node is an Inquired-Node



FIGURE II. IF INCEPTOR-NODE IS INQUIRED-NODE

As Figure 2 shows, Inceptor-Node *C* is an Inquired-Node. After node *B* eliminated, the topological structure is the right figure in Figure 2. Calculate the probability distribution as follows:

$$\begin{aligned} P(C) &= \sum_A \sum_E P'(A, C, E) \\ &= \sum_A \sum_E P(A) P(E) P'(C | A, E) \\ &= \sum_A \sum_E \sum_B P(A) P(B | A) P(C | B, E) P(E) \\ &= \sum_A \sum_E P(A) P(E) \sum_B P(B | A) P(C | B, E) \end{aligned}$$

$$P'(C | A, E) = \sum_B P(B | A) P(C | B, E), \quad P'(A | C) = P(A | C).$$

3) If the Inceptor-Node is an Ordinary-Node

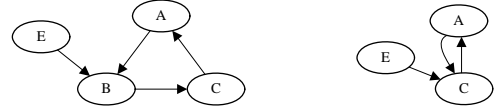


FIGURE III. IF INCEPTOR-NODE IS ORDINARY-NODE

As Figure 3 shows, Inceptor-Node *B* is an Ordinary-Node. After node *B* is eliminate, the topological structure is the right figure in Figure 3. Calculate the probability distribution as follows:

$$\begin{aligned} P(C) &= \sum_A \sum_E P'(A, E, C) \\ &= \sum_A \sum_E P(A) P(E) P'(C | A, E) \\ &= \sum_A \sum_E \sum_B P(A) P(E) P(B | A, E) P(C | B) \\ &= \sum_A \sum_E P(A) P(E) \sum_B P(B | A, E) P(C | B) \\ P'(C | A, E) &= \sum_B P(B | A, E) P(C | B). \end{aligned}$$

Synthesize the above three instances, Lemma 2 is proofed.

**Theorem 1** Any directed cycle can be simplified to a Simple-Cycle via Node-Elimination operation.

Lemma 1 indicates that no-inputs cycle can be simplified to a Simple-Cycle, and Lemma 2 indicates the cycle possessing inputs can also be simplified to a Simple-Cycle, so Theorem 2 is tenable. That is to say, every directed cycle on a cyclic BN can be simplified to a Simple-Cycle.

III. CONVERGENCY ANALYSIS FOR DIRECTED CYCLE

**Lemma 3**<sup>[7]</sup> If every node on a Simple-Cycle has just two values, then the cycle's probability distribution is convergent.

Proof: The Simple-Cycle's structure is showed in Figure 4, node *A* and *B* are both two-valued nodes,  $A(A_1, A_2)$ ,  $B(B_1, B_2)$ .



FIGURE IV. SIMPLE CYCLE

Let  $P(A_i | B_j) = \alpha_{ij}$ ,  $P(B_i | A_j) = \beta_{ij}$  where  $i = 1, 2$  and  $j = 1, 2$ , and the initial probability state is  $B(B_1 : P_{b1}^0, B_2 : P_{b2}^0)$ .

1) After the first propagation, probability states of each node are  $A(A_1 : P_{a1}^1, A_2 : P_{a2}^1)$  and  $B(B_1 : P_{b1}^1, B_2 : P_{b2}^1)$ :

$$P_{a1}^1 = P_{b1}^0 \alpha_{11} + P_{b2}^0 \alpha_{21},$$

$$P_{a2}^1 = P_{b1}^0 \alpha_{12} + P_{b2}^0 \alpha_{22},$$

$$\begin{aligned}
 P_{b1}^1 &= P_{a1}^0 \beta_{11} + P_{a2}^0 \beta_{21} = (P_{b1}^0 \alpha_{11} + P_{b2}^0 \alpha_{21}) \beta_{11} + (P_{b1}^0 \alpha_{12} + P_{b2}^0 \alpha_{22}) \beta_{21} \\
 &= P_{b1}^0 (\alpha_{11} \beta_{11} + \alpha_{12} \beta_{21}) + P_{b2}^0 (\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}) \\
 &= (\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}) + ((\alpha_{11} \beta_{11} + \alpha_{12} \beta_{21}) - (\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21})) P_{b1}^0 \\
 &\stackrel{\Delta}{=} w + s P_{b1}^0
 \end{aligned}$$

$$P_{b2}^1 = P_{a1}^1 \beta_{12} + P_{a2}^1 \beta_{22},$$

where  $\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} \stackrel{\Delta}{=} w$ ,

$$(\alpha_{11} \beta_{11} + \alpha_{12} \beta_{21}) - (\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}) \stackrel{\Delta}{=} s,$$

2) After probability propagate  $n$  times, the probability distribution of node  $B$ 's state  $B_1$  is:

$$\begin{aligned}
 P_{b1}^n &= w + s P_{b1}^{(n-1)} \\
 &= w + s(w + s P_{b1}^{(n-2)}) \\
 &= w(1 + s) + s^2 P_{b1}^{(n-2)} \\
 &= w(1 + s + \dots + s^{n-1}) + s^n P_{b1}^0, \text{ and}
 \end{aligned}$$

$$\begin{aligned}
 P_{a1}^n &= P_{b1}^{n-1} \alpha_{11} + P_{b2}^{n-1} \alpha_{21} \\
 &= \alpha_{21} + (\alpha_{11} - \alpha_{21}) [w(1 + s + \dots + s^{n-2}) + s^{n-1} P_{b1}^0],
 \end{aligned}$$

where

$$\begin{cases} w = \alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} \leq 1 \\ s = (\alpha_{11} \beta_{11} + \alpha_{12} \beta_{21}) - (\alpha_{21} \beta_{11} + \alpha_{22} \beta_{21}), |s| \leq 1 \end{cases}$$

If  $|s| < 1$ , Then:  $\lim_{n \rightarrow \infty} (1 + s + \dots + s^n) = \frac{1}{1-s}$ ,  $\lim_{n \rightarrow \infty} s^n = 0$ .

So

$$P_{b1}^* = \lim_{n \rightarrow \infty} P_{b1}^n = \frac{w}{1-s} \tag{2}$$

$$P_{a1}^* = \lim_{n \rightarrow \infty} P_{a1}^n = \alpha_{21} + (\alpha_{11} - \alpha_{21}) \frac{w}{1-s} \tag{3}$$

If  $s = 1$ , then  $\begin{cases} \alpha_{11} = \beta_{11} = 1 \\ \alpha_{21} = \beta_{21} = 0 \end{cases}$ , or  $\begin{cases} \alpha_{12} = \beta_{12} = 1 \\ \alpha_{22} = \beta_{22} = 0 \end{cases}$ . Well then:

$$w = \alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} = 0, \text{ and}$$

$$\begin{cases} P_{b1}^n = w(1 + s + \dots + s^{n-1}) + s^n P_{b1}^0 = P_{b1}^0 \\ P_{a1}^n = \alpha_{21} + (\alpha_{11} - \alpha_{21}) [w(1 + s + \dots + s^{n-2}) + s^{n-1} P_{b1}^0] = \alpha_{21} P_{b2}^0 + \alpha_{11} P_{b1}^0 \end{cases}$$

If  $s = -1$ , then  $\begin{cases} \alpha_{11} = \beta_{11} = 1 \\ \alpha_{21} = \beta_{21} = 0 \end{cases}$ , or  $\begin{cases} \alpha_{22} = \beta_{21} = 1 \\ \alpha_{12} = \beta_{11} = 0 \end{cases}$ .

So, when  $\begin{cases} \alpha_{21} = \beta_{11} = 1 \\ \alpha_{22} = \beta_{21} = 0 \end{cases}$ ,  $w = \alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} = 1$ :

$$P_{b1}^n = w(1 + s + \dots + s^{n-1}) + s^n P_{b1}^0 = \begin{cases} P_{b1}^0 & n \text{ is even} \\ P_{b2}^0 & n \text{ is odd} \end{cases},$$

$$P_{a1}^n = \alpha_{21} + (\alpha_{11} - \alpha_{21}) [w(1 + s + \dots + s^{n-2}) + s^{n-1} P_{b1}^0] = \begin{cases} P_{b1}^0 & n \text{ is odd} \\ P_{b2}^0 & n \text{ is even} \end{cases}$$

The probability distribution of each node on the cycle is

$$\begin{cases} A(A_1, A_2) : (P_{b1}^0, P_{b2}^0), (P_{b2}^0, P_{b1}^0), (P_{b1}^0, P_{b2}^0), (P_{b2}^0, P_{b1}^0) \dots \\ B(B_1, B_2) : (P_{b2}^0, P_{b1}^0), (P_{b1}^0, P_{b2}^0), (P_{b2}^0, P_{b1}^0), (P_{b1}^0, P_{b2}^0) \dots \end{cases}$$

When  $\begin{cases} \alpha_{22} = \beta_{21} = 1 \\ \alpha_{12} = \beta_{11} = 0 \end{cases}$ , then  $w = \alpha_{21} \beta_{11} + \alpha_{22} \beta_{21} = 1$ :

$$P_{b1}^n = w(1 + s + \dots + s^{n-1}) + s^n P_{b1}^0 = \begin{cases} P_{b1}^0 & n \text{ is even} \\ P_{b2}^0 & n \text{ is odd} \end{cases},$$

$$P_{a1}^n = \alpha_{21} + (\alpha_{11} - \alpha_{21}) [w(1 + s + \dots + s^{n-2}) + s^{n-1} P_{b1}^0] = \begin{cases} \alpha_{11} P_{b1}^0 + \alpha_{21} P_{b2}^0 & n \text{ is odd} \\ \alpha_{11} P_{b2}^0 + \alpha_{21} P_{b1}^0 & n \text{ is even} \end{cases}.$$

The probability distribution of each node on the cycle is:

$$\begin{cases} A(A_1, A_2) : (\alpha_{11} P_{b1}^0 + \alpha_{21} P_{b2}^0, \alpha_{11} P_{b2}^0 + \alpha_{21} P_{b1}^0), (\alpha_{11} P_{b2}^0 + \alpha_{21} P_{b1}^0, \alpha_{11} P_{b1}^0 + \alpha_{21} P_{b2}^0) \dots \\ B(B_1, B_2) : (P_{b2}^0, P_{b1}^0), (P_{b1}^0, P_{b2}^0), (P_{b2}^0, P_{b1}^0), (P_{b1}^0, P_{b2}^0) \dots \end{cases}$$

Synthesize all the above discusses, when  $-1 < s \leq 1$ , the probability distribution of every node on the cycle is convergent; when  $s = -1$ , the probability distribution is known. We call the phenomena that every node's probability distribution is constringent as **probability convergency** of the directed cycle.

**Lemma 4** Simple-Cycle possesses probability convergency.

Proof: As Fig.4 shows, the probability distribution of each node is  $P(A|B)$ ,  $P(B|A)$ .

1) if node  $B$  is multi-valued  $(b_1, b_2, \dots, b_n), n > 2$ , and

node  $A$  is two-valued, then:

Taken  $b_1$  as a dummy state  $B_1$ , and  $(b_2, \dots, b_n)$  as another dummy state  $B_2$ , calculate the new probability distribution for each dummy state:  
 $P'(B|A) = (P(B_1|A), P(B_2|A))$  and  
 $P'(A|B) = (P(A|B = B_1), P(A|B = B_2))$ :

$$P(B_2|A) = \sum_{i=2}^n P(B = b_i | A), P(A|B_2) = \sum_{i=1, n/2+1}^n P(b_i) P(A|b_i).$$

Thus there are only two nodes on the cycle, and each node is two-valued, then the cycle possesses probability constringency due to Lemma 3. Let

$(P^*(a_1), P^*(a_2), P^*(B_1), P^*(B_2))$  be the ultimate probability distribution, then:

$$P^*(b_1) = P^*(B_1)$$

$$P^*(b_i) = \sum_A P^*(A)P(b_i | A) = P^*(a_1)P(b_i | a_1) + P^*(a_2)P(b_i | a_2) \quad i = 2, \dots, n \quad (4)$$

2) if node  $A$  and node  $B$  are both multi-valued, then let  $A$  be  $\{a_m\}, m > 2$  and  $B$  be  $\{b_n\}, n > 2$  in a general way. So:

Taken  $a_1$  as a dummy state  $A_1$ , and  $(a_2, \dots, a_n)$  as another dummy state  $A_2$ , then we know that the cycle possesses probability constringency as we just discussed in 1). Let  $P^*(A_1), P^*(A_2), P^*(b_1), \dots, P^*(b_n)$  be the ultimate probability distribution, where

$$P^*(b_i) = \sum_A P^*(A)P(b_i | A) = P^*(A_1)P(b_i | A_1) + P^*(A_2)P(b_i | A_2) \quad i = 1, \dots, n \quad (5)$$

$$P^*(a_j) = \sum_{i=1}^n P^*(b_i) \times P(a_j | b_i) \quad j = 1, \dots, m \quad (6)$$

So, the ultimate probability state of the cycle is  $(P^*(a_j), P^*(b_j)) \quad i = 1, \dots, m; \quad j = 1, \dots, n$ .

Synthesize the discusses of 1) and 2), Lemma 4 is proofed.

**Theorem 2** A directed cycle's local probability distribution on a cyclic BN is convergent.

Theorem 2 can be educed easily because Theorem 1, Lemma3 and Lemma4 were proofed.

Any cycle on a cyclic BN must converge to some stable probability distribution  $P^*$ , and  $P^*$  can be calculated by Equation (2)-(6).

#### IV. PROBABILITY PROPAGATION FOR CYCLIC BAYESIAN NETWORK

Directed cycle's probability convergency is the theoretic base for probability propagation. Let probability distribution  $P^*$  is a directed cycle's convergent state, then the cycle will be treated as a special node of the BN with probability distribution  $P^*$ , so all the probability inference method introduced in Reference 3 can be used to update belief.

But in generally, engineering practice can not repeat endless. So, switch node is introduced to solve probability propagation problem whose position and value  $\varepsilon$  is determined by engineering practice. Assuming the local probability distribution of the cycle is  $P^*$  when the switch node approach to  $\varepsilon$ , cut the fan-in arc or fan-out arc to destroy the cycle. It is important for software project engineering practice to destroy the cycle when local probability reaches some point, because the ultimate purpose is to reach some satisfied situation within limited budget not to perfection. For the extreme instance  $s = -1$  as discussed in section 3,

introducing switch node is the only probability inference method.

A simple example is given in Figure 5 to show how the probability distribution propagate in the cycle. Software test is an iterative process usually, expressed as figure I in Figure 5. Conditional probability of each node is known, node  $A$  is evidence node, and  $B$ 's initial state is known also.

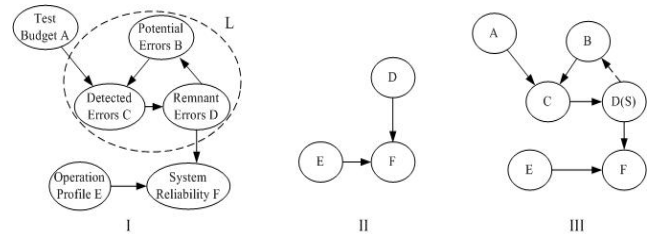


FIGURE V. PROBABILITY INFERENCE OF CYCLIC BAYESIAN NETWORK

Given the cycle  $L$ 's joint probability distribution converging to  $(P^*(B), P^*(C), P^*(D))$ , then as figure II shows, taking node  $D$  as evidence node, carry on probability propagation in the all the network via Equation (1).

In software project practice, when the remnant errors less than  $\varepsilon$ , system test will always be stopped. So, we choose node  $D$  as switch node and its threshold value is  $\varepsilon$ . When  $P^{k-1}(D < \varepsilon) < P^k(D < \varepsilon) > P^{k+1}(D < \varepsilon)$ , evaluate local joint probability distribution as  $P^* = P^k$ , and disconnect the directed arc  $D \rightarrow B$ . Then as figure III shows, we can calculate the joint probability distribution for all nodes in the network according to Equation (1).

#### V. CONCLUSION

Modeling the uncertainties in the software project process by Bayesian Network, provides a useful method to identify, evaluate and reduce latent risks in software project process. Modeling the iterative development of the software project based on Bayesain Network, results in the directed cycles in the BN process model. BN is a graphical model that encodes relationships among variables of interest, and it is a directed acyclic graph which has no ability to deal with the directed cycles. In this paper, **Theorem 1** indicates that, all the ordinary nodes on the directed cycle can be eliminated one by one until the cycle is converted into a simple cycle, and **Theorem 2** indicates that any simple cycle possesses probability astringency, so all the directed cycles are probability convergent. Based on probability convergency of the directed cycle, two kinds of probability propagation route are proposed.

In this paper, we discuss cyclic BN premising the nodes are all discrete variables, how about the continuous variables? We will discuss further in the future.

#### REFERENCE

[1] LI Ming-Shu, YANG Qiu-Song, ZHAI Jian. Systmatic Review of Software Process Modeling and Analysis[J]. Journal of Software, March 2009: 524-545. in Chinese  
[2] Bayesian Belief Network and Its Applications, Technical Report UVA-CSCS-BBN-001 Draft, May 20, 2002.

- [3] Haipeng Guo, William Hsu. A Survey of Algorithms for Real-Time Bayesian Network Inference. AAAI/KDD/UAI-2002 Joint Workshop on Real-Time Decision Support and Diagnosis Systems, Edmonton, July 2002.
- [4] WANG Jun, ZHOU Wei-Da. Research and Progress for Bayesian Networks[J]. Electronic Science and Technology, August 1999. in Chinese
- [5] Martin Neil, Norman Fenton. Predicting Software Quality using Bayesian Belief Networks. Proceedings of 21st Annual Software Engineering Workshop NASA/Goddard Space Flight Centre, 1996.
- [6] Norman Fenton, Martin Neil. Software Metrics and Risk[C]. 2nd European Software Measurement Conference, 8 October, 1999.
- [7] HU Yu-Sheng. Research and Application for dynamic bayesian network[D]. PHD for Beijing university of science and technology, May 2001. in Chinese
- [8] David Heckerman, a Tutorial on Learning with Bayesian Networks, Technical Report, MSR-TR-95-06, Microsoft Research, 1996.
- [9] Van der Gaag. Bayesian Belief networks: Odds and Ends [J], the Computer Journal, 39:97-113, 1996.