

Incomplete Information and Bayesian Knowledge-Bases

Eugene Santos, Jr. and Qi Gu

Thayer School of Engineering

Dartmouth College

Hanover, NH., USA

{Eugene.Santos.Jr, Qi.Gu}@Dartmouth.edu

Eunice E. Santos

Department of Computer Science

The University of Texas at El Paso

El Paso, TX, USA

eesantos@utep.edu

Abstract— Knowledge acquisition is an essential process in improving the problem-solving capabilities of existing knowledge-based systems through the absorption of new information and facilitating change in current knowledge. However, without a verification mechanism, these changes could result in violations of semantic soundness of the knowledge causing inconsistencies and ultimately, contradictions. Therefore, maintaining semantic consistency is of primary concern, especially when dealing with incompleteness and uncertainty. In this paper, we consider the semantic completability of a knowledge system as a means of ensuring long-term semantic soundness. In particular, we focus on how to preserve semantic completability as the knowledge evolves over time. Among numerous methods of knowledge representation under uncertainty, we examine Bayesian Knowledge-Bases, which are a rule-based probabilistic model that allows for incompleteness and cycles between variables. A formal definition of full/partial completability of BKB is first introduced. A principle to check the overall completability of a BKB is then formulated with a formal proof of correctness. Furthermore, we show how to use this principle as a guide for maintaining semantic soundness and completability during incremental knowledge acquisition. In particular, we consider two primary modifications to the knowledge base: 1) adding/fusing knowledge, and 2) changing/tuning conditional probabilities.

Keywords- semantic completability; incomplete information; knowledge representation; Bayesian knowledge-bases

I. INTRODUCTION

The development of a knowledge-based system is fundamentally an incremental process. The domain knowledge engineers (subject matter experts) are inclined to start by selecting a subarea that they personally understand well, and then continuously introduce new information to the existing knowledge-based system as a basic cyclical process [1], [2]. However, significant changes may take place at any step of the process, anywhere in the system, such as deleting or modifying existing knowledge or adding new knowledge. Any of these could introduce a contradiction to the existing semantics or be forced to lose some important prior information relationships. Considering that part of the knowledge may have been fully specified by knowledge engineers, losing such information could compromise the correctness of the overall knowledge base. As a simple example, if we raise the probability of “Rain” without decreasing the chance of “Not Rain”, we will end up

with a violation of probability theory. Therefore, one key issue during knowledge acquisition is to ensure semantic consistency. This is especially important to knowledge engineers who treat consistency as a basic rule of thumb while potentially using a generate-and-test means of constructing knowledge-bases.

On the other hand, the knowledge-based system at any stage is necessarily associated with some degree of incompleteness and uncertainty. This results in a natural expectation that a knowledge-based system should exhibit enough flexibility and intuitiveness for capturing knowledge in order to make it easy and responsive for the knowledge engineers to build models/systems that they can still understand, maintain, validate, update, and so forth. Bayesian approaches have been widely used for managing uncertainty. Among those, Bayesian Networks (BNs) [3], [4] are a popular probabilistic model due to its sound theoretical foundations in probability theory combined with efficient reasoning. Similar techniques generated from BN variants include first-order Bayesian logic [5] and semi-qualitative Probabilistic network [6]. One advantage of BNs comes from the ability to preserve semantic completability by nature, as they require a completely specified (filled) conditional probability table (CPT), which guarantees that the probabilities of all possible outcomes of the world sum up to 1. However, when the knowledge available on parameters is vague or incomplete, those restrictions turn out to be problematical since this becomes a major source of inflexibility. In addition, the causal mechanism in reality is likely to vary across the instantiations, which happens when the causal link between two variables becomes weak given some evidence. Under this situation, the underlying causal relationship may end up with a cycle, which is not permitted in a BN.

To avoid these limitations, we represent our knowledge using Bayesian Knowledge-Bases (BKBs) [7], [8]. BKBs are a rule-based probabilistic model that represents possible world states and their causal relationship using a directed graph. BKBs subsume BNs by specifying dependence at the instantiation level (versus BNs that specify only at the random variable level); by allowing for cycles between variables; and, by loosening the requirement in specifying probability distributions, and thus allowing for incompleteness. There are some other techniques that attempt to handle incompleteness, such as Fuzzy representations [9], imprecise probabilities

representations [10], and first-order language based knowledge representations [11]. However, they all suffer from potentially weak semantics for complex reasoning chains under uncertainty.

Though reasoning over incomplete information can be approximated through imprecise assessments [18], [19], without a verification mechanism with regards to semantic consistency, reasoning explanations are not sound. Given that BKBs focus on managing incomplete information, our goal in this paper is to further address the ability of BKBs on preserving complete semantics during the process of knowledge acquisition. In particular, we introduce formal conditions concerning the semantic completeness of a BKB and prove theorem correctness. In our previous work [8], theoretical and algorithmic results have been presented on how BKBs can implicitly preserve conditional semantics as a knowledge-base evolves. However, that work is based on the assumption of semantic completeness, whereas in this paper, we show what conditions are sufficient to achieve semantic completeness. Moreover, we discuss how semantic completeness can be maintained during incremental knowledge acquisition. Two types of modification are frequently used in knowledge base updating. One is knowledge fusion which integrates knowledge from different sources into one overall system. The other is system tuning that focuses on correcting a knowledge-base with minimal global change. We show how the conditions that guarantee the semantic completeness can be easily applied to the two algorithms without increasing computational complexity.

We begin in the next section with a detailed description of BKBs. In Section 3, we present a formal definition of probabilistically completeness and sufficient conditions for evaluating completeness, followed by additional properties of probabilistically completeness. Applications on two types of system modifications will be introduced in the next section. Finally, we will present our conclusions and future work.

II. BAYESIAN KNOWLEDGE BASES

BKBs subsume BNs by providing a more compact knowledge representation. Instead of specifying the causal structure using conditional probability tables (CPT) as in BNs, BKBs collect the conditional probability rules (CPR) in an “if-then” style. Fig. 1 shows a graph structure of a BKB fragment, in which A is a random variable with possible instantiations $\{a_1, a_2\}$. Each instantiation of a random variable is represented by an I-node, or “instantiation node”, e.g. $A = a_1$, and the rule specifying the conditional probability of an I-node is encoded in an S-node, or “support node” with a certain weight/probability. For example, q_8 corresponds to a CPR which can be interpreted as: if $A = a_2$ and $C = c_1$, then $B = b_2$ with a probability 0.1.

Rule-based models like BKBs have significant advantage in size of the representation since they allow for independence to be specified at the instantiation level and do not require the full table representation of the CPTs. This design endows BKBs with the capability of handling incomplete probabilistic knowledge. As the BKB shows in Fig.1, the dependency relationship at the variable level implies that variable B

depends on both A and C . However, given evidence $A = a_1$, B becomes independent of C . This could happen in the real world when the role of a critical variable can dominate some local dependency relationships between variables. Under this circumstance, there is no need to fill in $P(B|A = a_1, C)$ since BKBs only capture the knowledge that is available and does not require a complete distribution specification.

The other feature of BKBs is that they also allow cyclic relationships among random variables. Imagine if the direction of some causal mechanism also depends on specific states of the variables. Santos et al. [12] gives an example of BKB modeling of a political election, in which the type of “race” may flip the causal direction between the belief of a piece of evidence and the voting action.

The formal definition of the graphical representation of a BKB from Santos & Santos [7] is given below:

Definition 1. A *correlation-graph* is a directed graph $G = (I \cup S, E)$ in which $I \cap S = \emptyset, E \subset \{I \times S\} \cup \{S \times I\}$, and $\forall q \in S$, there exists a unique $a \in I$ such that $(q, a) \in E$. If there is a link from $q \in S$ to $a \in I$, we say that q supports a .

For each S-node q in a correlation graph G , we denote $Pred_G(q)$ as the set of I-nodes pointing to q , i.e. $Pred_G(q) = \{a \in I | a \rightarrow q \in E\}$ and $Desc_G(q)$ as the I-node supported by q , i.e. the a such that $q \rightarrow a \in E$.

Two I-nodes, α_1 and α_2 are said to be *mutually exclusive* if they are different instantiations of the same random variable. Similarly, two sets of I-nodes I_1 and I_2 are mutually exclusive if there exists two I-nodes $a_1 \in I_1$ and $a_2 \in I_2$, such that a_1 and a_2 are mutually exclusive. For example, the sets of I-nodes $\{A = a_1, B = b_2\}$ and $\{A = a_2, B = b_2, C = c_1\}$ are mutually exclusive.

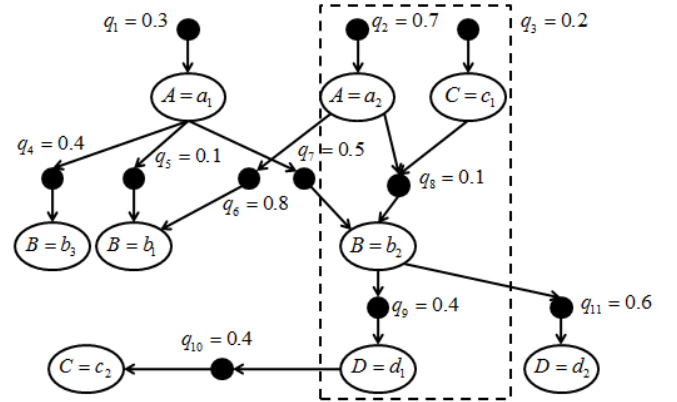


Figure 1: Example of a BKB fragment

Definition 2. A *state* θ is a set of I-nodes such that it contains no more than one instantiation of each random variable. θ is said to be *complete w.r.t* a set of random variables T if it contains exactly one I-node of variable in T .

Definition 3. A set of S-nodes R is said to be *complementary* if for all $q_1, q_2 \in R$, $Desc_G(q_1)$ and $Desc_G(q_2)$ are mutually exclusive, but $Pred_G(q_1)$ and $Pred_G(q_2)$ are not mutually exclusive. Variable v is said to be the *consequent variable* of

R , if for any S-node $q \in R$, $Desc_G(q)$ is an instantiation of v . R is said to be *maximum complementary* if there is no R' such that $R \subseteq R'$.

Notation: Let ρ_v denote the set that contains all possible complementary sets of S-nodes w.r.t. variable v , such that for any complementary set $r \in \rho_v$, v is the consequent variable of r . We also introduce Ψ_v to denote the subset of ρ_v that only consists of maximum complementary sets, i.e. $\Psi_v = \{r | r \in \rho_v \wedge \nexists r' \in \rho_v, r \subseteq r'\}$. For example in Fig. 1, $\{q_5, q_7\}$ is a complementary set w.r.t. variable B and $\Psi_B = \{\{q_4, q_5, q_7\}, \{q_6, q_8\}\}$.

Definition 4. A BKB is a tuple $K = (G, w)$ where $G = (I \cup S, E)$ is a correlation-graph, and $w : S \rightarrow [0, 1]$ such that

1. $\forall q \in S, Pred_G(q)$ contains at most one instantiation of each random variable.
2. For distinct S-nodes $q_1, q_2 \in S$ that support the same I-node, $Pred_G(q_1)$ and $Pred_G(q_2)$ are mutually exclusive.
3. For any complementary set of S-nodes $R \subseteq S$, R is normalized: $\sum_{q \in R} w(q) \leq 1$ where $w(q)$ is a weight function that represents the conditional probability $P(Desc_G(q) | Pred_G(q))$.

The intuition behind these three conditions is that each S-node can only support one I-node; two rules supporting the same I-node cannot be satisfied at the same time; and, to ensure normalization of the probability distribution, every complementary set of S-nodes should be normalized. Given a variable v , since any complementary set in ρ_v is a subset of some complementary set in Ψ_v , the last condition can be translated into: for any variable v , $R \in \Psi_v$ should be normalized.

Like BNs, reasoning with BKB is also based on the calculation of joint probabilities over the possible states, which can be captured as a subgraph of a BKB, called an *inference*.

Definition 5. Let $K = (G, w)$ be a BKB with correlation graph $G = (I \cup S, E)$. A subgraph $\tau = (I' \cup S', E')$ of G is called an *inference* over K if

1. τ is acyclic.
2. (*Well-supported*) $\forall a \in I', \exists q \in S', q \rightarrow a \in E'$
3. (*Well-founded*) $\forall q \in S', Pred_\tau(q) = Pred_G(q)$
4. (*Well-defined*) $\forall q \in S', Desc_\tau(q) = Desc_G(q)$
5. There is at most one I-node corresponding to any given random variable in I' . Furthermore, τ is said to be a *complete inference* over K if state I' is complete w.r.t. all random variables.

For example in Fig.1, the dotted rectangle encloses a complete inference which has a complete state $\{A = a_2, B = b_2, C = c_1, D = d_1\}$. The joint probability of an inference τ is just the product of the weights of all S-nodes in τ , i.e. $P(\tau) = \prod_{q \in \tau} w(q)$. The idea of inference plays an important role in two forms of reasoning with BKBs, belief revision (also called maximum a posteriori or MAP) [3], [7], and belief updating, where belief revision tries to determine the most probable state given some evidence and belief updating tries to calculate the marginal probability of a partial state. BKBs have been applied to a variety of real-world problems

such as adversary intent modeling [13], social events prediction and analysis [14], and surgical intent modeling [15].

III. COMPLETABILITY ANALYSIS

BKBs are by nature designed to handle incomplete information. However, as we mentioned earlier, it is critical to preserve the existing semantic completability as new knowledge is incrementally added to the knowledge-base. Hence, a verification mechanism is necessary to check/preserve semantic completability [16], [17]. In this section, we show how to preserve semantic completability in BKBs by first examining the sufficient conditions for guaranteeing overall semantic completability. Unlike BNs that model uncertainties using complete specifications of conditional probability distributions, BKBs do not have these restrictions in the first place, which increases the flexibility for knowledge representation. Furthermore, there are many situations during knowledge acquisition when the conditional probability distribution cannot be fully specified, such as when the knowledge engineers are not fully confident about the whole semantic space or the information regarding model parameters is simply unavailable. Though such limitations may lead to an incomplete overall semantics, it is still helpful to maintain the partial completability of such variables that are conditionally independent of those loosely specified variables. For example, suppose that “rain” has a direct effect on the “sprinkler level” (when it rains, the level of sprinkler is usually low), then even if we lack statistics on the conditional probability of “sprinkler level” given “rain”, it is still useful to make sure that the probabilities of “rain” and “not rain” sum up to 1. Thus, our second task is to provide conditions for achieving partial completeness. Lastly, properties entailed by the semantic completability will be formally derived.

A. Completeness

Given a BKB $K = (G, w)$, let $T = \{X_1, X_2, \dots, X_n\}$ be a set of random variables in K and I_T be the set containing all complete states corresponding to T .

Key Definition 6. A random variable v is said to be *semantically complete* if $\sum_i P(v = i) = 1$. K is said to be *semantically complete* w.r.t. T if for any variable $v \in T$, v is semantically complete.

Key Definition 7. K is said to be *composite/partial complete* w.r.t. T if $\sum_{\theta \in I_T} P(\theta) = 1$. K becomes *fully complete* if T contains all random variables.

Definition 8. K is said to be *assignment complete* w.r.t. T if for any complete state $\theta \in T$, there exists a complete inference over K ; K is said to be *locally complete* w.r.t. T if for any variable $v \in T$, $\forall R \in \Psi_v, \sum_{q \in R} w(q) = 1$.

Lemma 1. For S-node q , if an inference τ contains both $Desc_G(q)$ and $Pred_G(q)$, then $q \in \tau$.

Proof. Assume that $q \notin \tau$, then to find an inference τ containing I-node $Desc_G(q)$, there must be another S-node q' supporting $Desc_G(q)$. However, all the S-nodes pointing to the same I-node must be mutual exclusive. In other words, q' and $Pred_G(q)$, cannot coexist. Contradiction. ■

Definition 9. Let $\theta = \{X_1 = x_1, X_2 = x_2, \dots, X_n = x_n\}$ be a state. An inference $\tau = (I \cup S, E)$ is said to be a *minimal inference w.r.t. θ* if $\theta \subseteq I$ and there does not exist any other inference $\tau' = (I' \cup S', E') \subseteq \tau$ such that $\theta \subseteq I'$.

Definition 10. Let q be the S-node supporting I-node α , α is said to be a *root* of K if $\text{Pred}_G(q) = \emptyset$.

Key Theorem 1: If K is both assignment complete and locally complete w.r.t. all random variables, K is fully complete.

Proof. We prove this by induction on the number of I-nodes in K . If K only has one I-node, namely $A = a$. From the definition of BKBs and since K is assignment complete, $A = a$ must have exactly one supporting S-node. Let q be this S-node. As K is locally complete, it's trivial to show that $\sum_{\theta \in I_T} P(\theta) = P(A = a) = 1$.

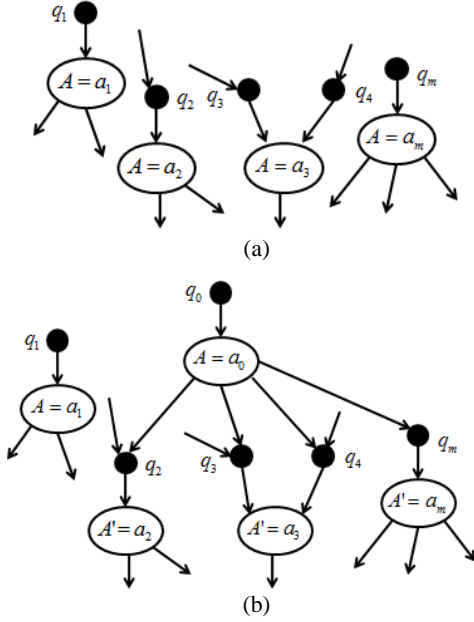


Figure 2: Example of transformation. (a) Original BKB K . (b) Transformed BKB K'

Assume that the theorem is true for all knowledge graphs up to $N - 1$ I-nodes. Let K be both an assignment and locally complete BKB with exactly N I-nodes. Let H be the set containing all complete inferences over K . Considering that every complete inference in H is also a minimum inference w.r.t. T (T corresponds to all variables in this case), the problem of proving $\sum_{\theta \in I_T} P(\theta) = 1$ is reduced to proving $\sum_{\tau \in H} w(\tau) = 1$ (Theorem 3.9 in Santos et al. [8]). Let $V(K)$ be the set of all roots of K . From Proposition 4.7 in Santos and Santos [7], $V(K)$ is not empty, e.g. I-node $a_1 \in V(K)$. Let $\sigma = \{a_1, a_2 \dots a_m\}$ be the set of I-nodes containing all instantiations of variable A .

If a_1 is the only instantiation of variable A , i.e. $|\sigma|=1$, let q_1 be its uniquely associated S-node. Then a_1 participates in all complete inferences in H and $w(q_1) = 1$. We remove a_1, q_1 , and all edges coming in and out of a_1 . Clearly, the remaining BKB is still assignment and locally complete, but with $N - 1$ I-nodes. By induction, $\sum_{\tau \in H} \prod_{q \in S_{\tau-q_1}} w(q) = 1$.

Thus, $\sum_{\tau \in H} w(\tau) = w(q_1) \sum_{\tau \in H} \prod_{q \in S_{\tau-q_1}} w(q) = 1$.

Otherwise, since $a_1 \in V(K)$, let X be the set of S-nodes found in K that support the remaining I-nodes $\{a_2, a_3 \dots a_m\}$. Now, construct a new $K' = (G', w')$ from K as follows:

1. Introduce new I-node a_0 and new S-node q_0 for a_0 . Let $w(q_0) = 1 - w(q_1)$.
2. Replace set σ by $\{a_1, a_0\}$ and introduce a new random variable A' with possible instantiations $\{a_2, a_3 \dots a_m\}$.
3. For each S-node $q \in X$, add edge (a_0, q) and change $w(q)$ to $w'(q) = w(q)/w(q_0)$.

Fig. 2 shows an example of transformation. From this construction, K' is still a valid BKB. Since the inferences which contain a_1 are not affected by the new construction, and for any inference τ that includes $q \in X$, clearly, $w'(\tau) = w(\tau)$. Thus, $\sum_{\tau' \in H'} w'(\tau') = \sum_{\tau \in H} w(\tau)$.

Next, we prove that K' is also locally complete. In fact, only the S-nodes associated with variable A and A' have been modified. It is trivial to prove locally completeness for A since $w(q_0) + w(q_1) = 1$. Let $R \in \Psi_{A'}$. In original BKB K , q_1 is complementary with any S-node $q \in R$, or $w(q_1) + \sum_{q \in R} w(q) = 1$ due to the fact that a_1 is a root node. Thus in new BKB K' , $w'(R) = \sum_{q \in R} w'(q) = \sum_{q \in R} w(q) / (1 - w(q_1)) = (1 - w(q_1)) / (1 - w(q_1)) = 1$.

Similar to K , let H' be the set containing all complete inferences in K' . We denote S_{τ} as the set of S-nodes corresponding to inference τ . We partition H' into two sets $\{H_1, H_0\}$ where

- $\tau \in H_0$ if and only if $a_0 \in \tau$ and
- $\tau \in H_1$ if and only if $a_1 \in \tau$.

Thus,

$$\begin{aligned} \sum_{\tau \in H'} w'(\tau) &= \sum_{\tau \in H_0} w'(\tau) + \sum_{\tau \in H_1} w'(\tau) \\ &= w'(q_1) \sum_{\tau \in H_1} \prod_{q \in S_{\tau-q_1}} w'(q) + w'(q_0) \sum_{\tau \in H_0} \prod_{q \in S_{\tau-q_0}} w'(q) \end{aligned}$$

Case 1.

Construct a new knowledge-base $K_1 = (G_1, w')$ from K' by combining all the inferences in H_1 into one subgraph G_1 , i.e. $G_1 = \bigcup_{\tau \in H_1} \tau$. Clearly, K_1 is a valid BKB, where a_1 is the only instantiation of variable A . We claim that:

1. K_1 is assignment complete.
2. K_1 is locally complete w.r.t. all variables but A .

The first claim is obvious. We prove the second one by contradiction. Assume there exists a variable v in K_1 , $v \neq A$, such that $\exists R \in \Psi_v, \sum_{q \in R} w'(q) < 1$. In other words, there used to be a S-node q^* in K' but not in K_1 that is also complementary with R . We denote $Pa(R)$ as the superset of all parent I-nodes in R , i.e. $Pa(R) = \bigcup_{q \in R} \text{Pred}_{G_1}(q)$. As the example shows in Fig. 1, for maximum complementary set of S-nodes $R = \{q_6, q_8\}$, $Pa(R) = \{(C = c_1), (A = a_2)\}$. From Lemma 1, $\text{Pred}_{G'}(q^*)$ must be mutually exclusive with a_1 , otherwise, q^* will not be precluded from K_1 . Let $a_i \in \text{Pred}_{G'}(q^*)$, $a_i \neq a_1$.

We claim that in K_1 , there exists another S-node pointing to $Desc_G(q^*)$ that is also complementary with R . The property of assignment complete guarantees that there is an inference τ that contains $\{Pa(R), a_1, Desc_G(q^*)\}$ in K_1 . Let q' be the S-node in τ , such that $(q', Desc_G(q^*)) \in E_1$, clearly, q' is complementary with R and $q' \neq q^*$. Thus, R is not maximum complementary. Contradiction.

Since K_1 is locally complete w.r.t. all variables but A , we can simply remove a_1, q_1 , and all edges coming in and out of a_1 . Then the remaining BKB is both assignment complete and locally complete with the number of I-nodes less than N . By induction, $\sum_{\tau \in H_1} \prod_{q \in S_\tau - q_1} w'(q) = 1$

Case 2.

Similar to Case 1, we construct a new knowledge-base K_0 from K' by combining all the inferences in H_0 into K_0 . Again, after removing q_0, a_0 and corresponding edges, the remaining BKB is fully complete with the number of I-nodes less than N . By induction, $\sum_{\tau \in H_0} \prod_{q \in S_\tau - q_0} w'(q) = 1$

Thus,

$$\begin{aligned} \sum_{\tau \in H'} w'(\tau) &= w'(q_1) \sum_{\tau \in H_1, q \in S_\tau - q_1} \prod w'(q) + w'(q_0) \sum_{\tau \in H_0, q \in S_\tau - q_0} \prod w'(q) \\ &= w'(q_1) + w'(q_0) = 1 \quad \blacksquare \end{aligned}$$

Key Theorem 2: If K is both assignment complete and locally complete w.r.t. T , K is *composite complete* w.r.t. T .

Sketch of Proof. Let H^* be the set containing all complete inferences corresponding to T . We construct a new BKB $K^* = (G^*, w)$, $G^* = \bigcup_{\tau \in H^*} \tau$. Clearly, K^* is assignment complete. The local completeness of K^* can be proved in a similar way as Theorem 1. Therefore, K^* is fully complete, and thus K is composite complete w.r.t. T .

Key Corollary 1, If K is both assignment complete and locally complete w.r.t. T , K is *semantically complete* w.r.t. T .

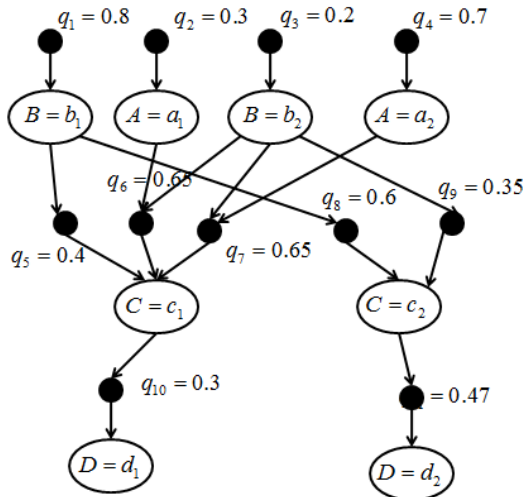


Figure 3: Example of a composite complete BKB

Theorem 2 and Corollary 1 can be applied to quantify the completability level of BKB in terms of the number of

semantically complete variables. We take the BKB K in Fig. 3 as an example. Though the conditional probability distribution of $P(D|C)$ is not fully specified, variables A, B and C still hold complete semantics. The complexity of checking for completability is the same as performing belief updating (i.e. computing posterior probabilities).

B. Properties for BKB completeness

As we proved in the last section, assignment complete and locally complete are two sufficient conditions for semantic completeness. However, are they also necessary? Moreover, can we derive other important BKBs properties from probabilistically completeness?

Note: If K is fully complete, K may not be assignment complete. Fig. 4 gives an example where $\sum_{i,j} P(A=i, B=j) = 1$, but there is no inference for $\theta_1 = \{A = a_1, B = b_3\}$ or $\theta_2 = \{A = a_2, B = b_1\}$. The reason is that the probabilities for states θ_1 and θ_2 are both equal to 0. Since BKBs only capture meaningful semantics, the redundant dependency relationships are removed.

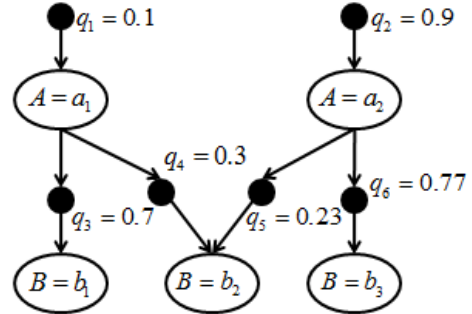


Figure 4: Example of a fully complete but not assignment complete BKB

Definition 11. A node $a \in I \cup S$ in BKB K is said to be *grounded* if there exists an inference τ over K such that a is in τ . K is said to be grounded if $\forall a \in I \cup S$, a is grounded.

Lemma 2. If K is fully complete and $\forall q \in S, w(q) > 0$, then K is grounded.

Proof. Assume there exists a S-node q that belongs to no inferences. Then from Lemma 1, there is no inference containing both $Desc_G(q)$ and $Pred_G(q)$ or $P(Desc_G(q), Pred_G(q)) = 0$. However, $w(q) = P(Desc_G(q)|Pred_G(q)) = P(Desc_G(q), Pred_G(q))/P(Pred_G(q)) > 0$, So, $P(Desc_G(q), Pred_G(q)) > 0$. Contradiction. \blacksquare

The underlying philosophy behind lemma 2 is that in a fully complete BKB, no rule is redundant. Thus, if a S-node remains grounded during knowledge acquisition, then its initial assigned probability can be semantically preserved [8].

Lemma 3. If K is fully complete and $\forall q \in S, w(q) > 0$, then K is locally complete.

Sketch of Proof. Assume there exists a variable v , such that $\exists R \in \Psi_v, \sum_{q \in R} w(q) < 1$. Then we can always select a S-node $q' \in R$ and increase $w(q')$ to $w'(q') = w(q') + 1 - \sum_{q \in R} w(q)$. From Lemma 2, q' must be grounded. Let τ' be an

inference with q' inside. Then $\sum_{\tau \in H} w'(\tau) = \sum_{\tau \in H - \tau'} w(\tau) + P(\tau') > 1$. Contradiction. ■

Lemma 3 guarantees that any change that affects the joint probabilities does not affect the local completeness as long as the semantic completability is preserved in the knowledge base.

IV. MAINTENANCE OF SEMANTIC CONSISTENCY DURING KNOWLEDGE ACQUISITION

The knowledge-based system allows knowledge engineers to potentially make significant changes to the knowledge base. Such changes, if not carefully handled, could lead to serious semantic conflicts and thus degrade the problem-solving capabilities of the system. In this section, we present how Theorems 1&2 can be readily applied to preserve semantic completeness with regards to two types of modification.

A. Application to BKB Fusion

One type of knowledge base modification occurs when several pieces of information are to be fused together. This happens frequently in decision making scenarios when multiple perspectives/experts are involved. For example, in the case of a medical diagnosis system, two experts may hold different opinions about the causal relationship between two variables of interest. So, simply combining the two knowledge bases could cause a violation of probabilistic soundness. An algorithm [12] was proposed to encode and aggregate a set of knowledge fragments from different sources into one unified BKB. An important property of this approach is their ability to support transparency in analysis. In other word, all perspectives are preserved in the fused BKB without losing any information. The idea behind this algorithm is to take the union of all input fragments by incorporating additional variables called *source nodes*, indicating the source and reliability of the fragments. An example of BKB fusion is shown in Fig. 5., where two fragments from two sources Dr. X and Dr. Y are fused into one, and the decision will be drawn based on which source of information is invoked. Furthermore, it was demonstrated that elements from different sources could also be combined to form new inferences.

Theorem 3. If all BKB fragments are semantically complete and all of the source nodes are locally complete, then the fused BKB is *semantically complete*.

Proof. From the fusion algorithm, we note that if two fragments i and j are sharing one I-node, say $A = a$, then after being fused together, only one fragment will be activated at one time due to the fact that two source node $S_A = i$ and $S_A = j$ are mutual exclusive. Let p be the distribution in the fusion of these two fragments, we have $p(S_A = i) + p(S_A = j) = 1$. Thus,

$$\begin{aligned} \sum_k p(A = k) &= \sum_k [p_i(A = k)p(S_A = i) + p_j(A = k)p(S_A = j)] \\ &= \sum_k p_i(A = k)p(S_A = i) + \sum_k p_j(A = k)p(S_A = j) \\ &= p(S_A = i) + p(S_A = j) \\ &= 1 \end{aligned}$$

where, p_i and p_j correspond to the original probability distribution from fragment i and j . ■

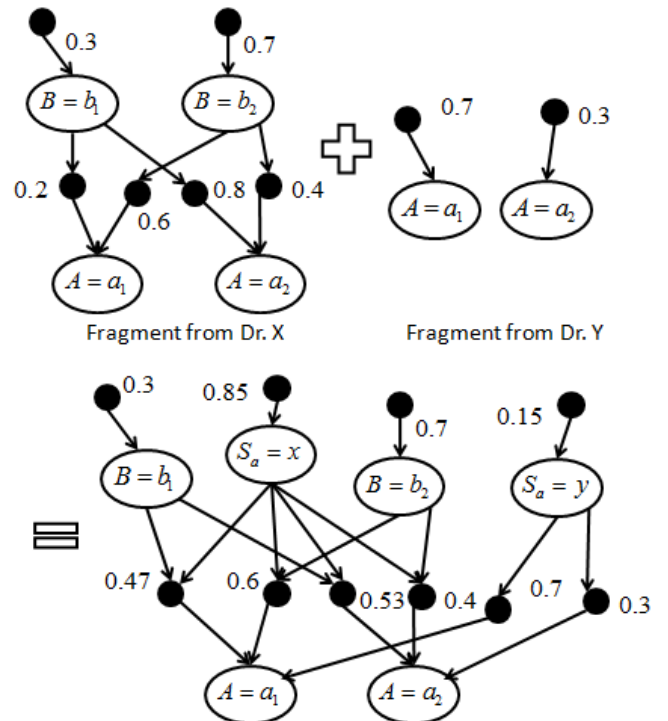


Figure 5. Example of BKB fusion of fragments from two different doctors.

Theorem 3 offers a sound theoretical foundation for preserving the semantic completeness in fragment knowledge during knowledge integration. As depicted in Fig. 5, the fused BKB from the bottom inherits the semantic completeness from the two fragments, e.g. $P(A = a_1) + P(A = a_2) = 1$. Moreover, Theorem 3 can be easily implemented in modeling real world scenarios since the only implementation required is to normalize the reliability of each source node.

B. Application to BKB tuning

Another type of modification takes the form of tuning the probabilities of certain rules. Santos et al. [20] proposed a method to correct a knowledge-based system by tuning necessary conditional probabilities while minimizing overall change in the knowledge-base. The basic idea is to transform the problem of multiple parameters changes as well as BKB structural constraints into a Linear Programming problem which can be solved efficiently. One of the constraints regulated by a BKB is that for any random variable v , $R \in \Psi_v$, R should be normalized, i.e. $\sum_{q \in R} w(q) \leq 1$. Considering that the tuning process only adjusts the weights of S-nodes without changing the structure of BKB, to preserve the completeness of the original BKB, we can simply replace the upper inequality with our locally complete constraint: $\sum_{q \in R} w(q) = 1$. Clearly, the new problem is still LP solvable.

In fact, Santos et al. [20] has shown that if the original BKB is locally complete, the tuned system also retains the property of locally complete. In other words, the best way to minimizing the global changes is to maintain the semantic completeness.

V. CONCLUSION

In this paper we proposed a verification mechanism to check the semantic completability of a knowledge-based system, such

that the initial value and semantics specified by the knowledge engineers can be preserved during incremental knowledge acquisition. We start by examining the sufficient conditions for guaranteeing overall semantic completeness and then extend to partial completeness. Furthermore, we derived some important properties concerning the completeness of BKBs. Another contribution of this work is that we demonstrate how this verification mechanism can be readily applied to preserve semantic completeness with regards to knowledge integration and system tuning without increasing computational complexity.

For future work, we will examine how the verification mechanism can be applied to knowledge validation for BKBs whose goal is to make the knowledge bases satisfy all test cases given by knowledge engineers. One type of knowledge validation is to deal with the problem of thrashing, where incompleteness is a main contributor to thrashing results. Therefore, to solve this problem, it becomes important to define and detect the incompleteness, such that the hinted information can be correctly added to the knowledge base and ultimately bridge the gap between the availability of amounts of information and expertise.

Acknowledgments. This research was supported in part by grants from DTRA, AFOSR, ONR, and DHS.

REFERENCES

- [1] A. J. Gonzalez and D. D. Dankel. "The Engineering of Knowledge-Based Systems: Theory and Practice. Prentice-Hall, Inc., 1993
- [2] R. Knauf, S. Tsuruta, A. J. Gonzalez, "Toward Reducing Human Involvement in Validation of Knowledge-Based Systems," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on , vol.37, no.1, pp.120-131, Jan. 2007
- [3] J. Pearl, "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", Morgan Kaufmann Publishers Inc., San Mateo, CA, 1988
- [4] T. Silander and P. Myllymäki. "A simple approach for finding the globally optimal Bayesian network structure". Proc. UAI, 2006
- [5] K. B. Laskey, "MEBN: A language for first-order Bayesian knowledge bases.", Artificial Intelligence 172(2-3):140-178, 2008
- [6] C.P. de Campos and F.G. Cozman, "Belief updating and learning in semi-qualitative probabilistic networks", Proc. UAI, pp. 153-160, 2005.
- [7] E. Santos, Jr., and E. S. Santos, "A framework for building knowledge bases under uncertainty". Journal of Experimental and Theoretical Artificial Intelligence. 11(2):265-286, 1999.
- [8] E. Santos, Jr., E. S. Santos, S. E. Shimony, "Implicitly preserving semantics during incremental knowledge base acquisition under uncertainty", in International Journal of Approximate Reasoning, vol. 33, issue 1, pp. 71-94, Apr. 2003
- [9] Mauris, G.P.. "Fuzzy representation of incomplete knowledge about infinite support probability distributions in a measurement context", Fuzzy Systems Conference, pp. 1-6, 2007
- [10] C. Baudrit and D. Dubois. "Practical Representations of incomplete probabilistic knowledge", Journal Computational Statistics & Data Analysis, vol. 51, Issue 1, 2006
- [11] R.P.A. Petrick and F. Bacchus, "Extending the Knowledge-Based Approach to Planning with Incomplete Information and Sensing.", in Proc. ICAPS, 2004
- [12] Santos, Jr. E., Wilkinson, J. T., and Santos, E. E. "Bayesian Knowledge Fusion". Proceedings of the 22nd International FLAIRS Conference. AAAI Press. 559-564, 2009
- [13] Pioch, N. J., Melhuish, J., Seidel, A., Santos, Jr. E., and Li, D. "Adversarial Intent Modeling Using Embedded Simulation and Temporal Bayesian Knowledge Bases", Proc of the SPIE Defense & Security Symposium, Vol. 7348, pp. 0E-1-0E-12, 2009
- [14] Santos, E. E., Santos, Jr. E., Wilkinson, J. T., and Xia, H. "On a Framework for the Prediction and Explanation of Changing Opinions". Proceedings of the IEEE International Conference on SMC, 1146-1452, 2009
- [15] Santos, Jr. E., Kim, K. J., Yu, F., Li, D., Jacob, E., Katona, L., and Rosen, J. "A method to improve team performance in the OR through intent inferencing". Proc. EUROSIM Congress on Modelling and Simulation Vol. 2 , 2010
- [16] Todd J. Green "Models for Incomplete and Probabilistic Information", Advances in Database Systems, Volume 35, 9-43, 2009
- [17] R. Bonichon and O. Hermant, "A semantic completeness proof for TaMed". LPAR, LNCS vol. 4246, pp. 167-181, 2006
- [18] Prade, Henri; , "A Computational Approach to Approximate and Plausible Reasoning with Applications to Expert Systems," Journal of Pattern Analysis and Machine Intelligence, vol.PAMI-7, no.3, pp.260-283, 1985
- [19] C. B. Anagnostopoulos , Y. Ntirladimas , S. Hadjiefthymiades, "Situational computing: An innovative architecture with imprecise reasoning", Journal of Systems and Software, v.80 n.12, p.1993-2014, 2007
- [20] E. Santos, Jr., Q. Gu and E. E. Santos. "Tuning a Bayesian Knowledge base", Proceedings of the 24th International FLAIRS Conference. AAAI Press, 2011