Bayesian Knowledge Fusion

Eugene Santos Jr. and **John T. Wilkinson**
Thayer School of Engineering
Dartmouth College
Hanover, NH 03755-8000

Eunice E. Santos
Department of Computer Science
Virginia Polytechnic Institute
Blacksburg, VA 24060

Abstract

We address the problem of information fusion in uncertain environments. Imagine there are multiple experts building probabilistic models of the same situation and we wish to aggregate the information they provide. There are several problems we may run into by naively merging the information from each. For example, the experts may disagree on the probability of a certain event or they may disagree on the direction of causility between two events (e.g., one thinks A causes B while another thinks B causes A). They may even disagree on the entire structure of dependencies among a set of variables in a probabilistic network. In our proposed solution to this problem, we represent the probabilistic models as Bayesian Knowledge Bases (BKBs) and propose an algorithm called *Bayesian knowledge fusion* that allows the fusion of multiple BKBs into a single BKB that retains the information from all input sources. This allows for easy aggregation and de-aggregation of information from multiple expert sources and facilitates multi-expert decision making by providing a framework in which all opinions can be preserved and reasoned over.

Introduction

There are many situations when information fusion can be useful, from fusion of low level sensor data to the aggregation of higher level models built by human experts. One simple example is the case of multiple medical diagnosis expert systems, each containing rules from different experts. One expert may believe that the causal link between disease A and symptom B is strong while another may think it is weak; or one may think that condition A causes condition B while another thinks that B causes A. In the former case we could take the max of the two conditional probabilities, or the average, or some other function of the two, but in doing so we would likely lose some information. In the latter case we would end up with a cyclic causal relationship, violating the rules of some knowledge representation schemes like the widely used Bayesian networks (Pearl 1988). The ability to aggregate knowledge from mutiple sources in light of these difficulties is crucial in decision making scenarios involving mutiple subject matter experts and perspectives.

In addition to combining the opinions of multiple experts, another major goal of this work is to allow the easy aggregation and de-aggregation of pieces of information. This allows for reasoning over only the relevant pieces of the knowledge base. It also alleviates the problem of the high computational cost of reasoning in very large knowledge bases and allows for focused source-based analyses. Ease of aggregation is also ideal when dealing with a probabilistic model of a dynamic or evolving situation where new knowledge is being gained and added to the knowledge base incrementally and old knowledge may be removed if it is no longer applicable.

In this paper, we use Bayesian Knowledge Bases (BKBs) (Santos and Santos 1999; Santos, Santos, and Shimony 2003; Rosen, Shimony, and Santos 2004) as our knowledge representation framework. A BKB is a generalization of a Bayesian network (BN) in which dependencies are specified between states of random variables instead of between random variables, directed cycles are allowed, and the probability distribution specified need not be complete. We show how to take a collection of BKBs and merge them into one single BKB. In the context of the fusion process, we refer to the input BKBs as Bayesian Knowledge Fragments (BKFs) since each represents one fragment of knowledge that we wish to fuse into a larger knowledge base. Thus each BKF contains one chunk of knowledge and the fusion process allows them to be composed into larger, semantically meaningful models. These fragments can be easily stored, retrieved, and aggregated.

In related work, others have tried to use BNs for knowledge representation, but techniques based on BNs run into problems when it comes to tractable manipulation and reasoning. These techniques fall victim to exponential run time complexity, the need for complete specification of probability distributions, and overly rigid rules regarding conditional dependence structure. Included in this class are BN variants such as object-oriented BNs (Koller and Pfeffer 1997), multi-entity BNs (Laskey 2008), and Bayesian blackboards (Sutton et al. 2004).

Our fusion algorithm is based on the addition of new nodes to the input fragments called *source nodes*. These nodes indicate which rules in the knowledge base came from which fragments. A *reliability index* is given to each fragment indicating the trustworthiness of the source of the

1. Clinton downplayed MLK = yes, Race = Black $\xrightarrow{0.75}$ Support = Obama
2. Race = White, Like Bill Clinton = yes, Hillary more electable = yes $\xrightarrow{0.80}$ Support = Clinton
3. Support = Obama $\xrightarrow{0.15}$ Campaign for Obama = yes
4. Support = Clinton $\xrightarrow{0.70}$ Clinton downplayed MLK = no

Figure 1: A BKB fragment from a knowledge base about the South Carolina Democratic Primary election as a set of CPRs.

knowledge contained in the particular fragment. With these source nodes and their corresponding reliabilities, the inference process on the fused BKB can consider information from multiple sources when constructing explanations for any evidence observed. Thus the constructed explanation may contain elements from several different sources.

In the next section we will describe BKBs in more detail. Then we will present the fusion algorithm. After that we will further investigate the properties of the fused BKB, and then we will end the paper with a conclusion and description of future directions for this research.

Bayesian Knowledge Bases

A BKB is a collection of conditional probability rules (CPRs) of the form *if* $A_1 = a_1, A_2 = a_2, \ldots, A_n = a_n$, *then* $B = b$ *with probability* p, that satisfies conditions of normalization and mutual exclusivity that will be formally defined shortly. In these rules, A_i and B are random variables and a_i and b are instantiations, or states, of those random variables. BKBs subsume Bayesian networks (BNs) as all BNs are representable as BKBs (one can form a BKB from a BN by making all the conditional probability table entries in the BN into CPRs in the BKB). In contrast to BNs, BKBs allow for independence to be specified between states of random variables instead of between the random variables themselves. They also do not require that the probability distribution specified by the rules be complete; inference is carried out on only those rules that are specified. This allows for reasoning in the face of incompleteness and alleviates the burden of filling in large conditional probability tables which can be especially problematic when all the probabilities for these tables are not readily available.

Similarly to a BN, a BKB can be represented graphically, but the graph is not required to be acyclic as with BNs. There are two types of nodes in a graphically depicted BKB, I-nodes representing the different instantiations of the random variables, and S-nodes, or "support nodes". An example of a BKB from work on the South Carolina Democratic Primary election is shown in rule form in Figure 1 and in graph form in Figure 2. Each solid black circle in Figure 2 is an S-node and corresponds to exactly one of the CPRs from Figure 1. The number next to the S-node is its weight. Each text-filled oval is an I-node corresponding to one state of one of the random variables. The dependencies shown in the BKB would result in a circular directed graph at the random variable level (Figure 3) and thus would not be a valid BN. The circular relationship results because black voters

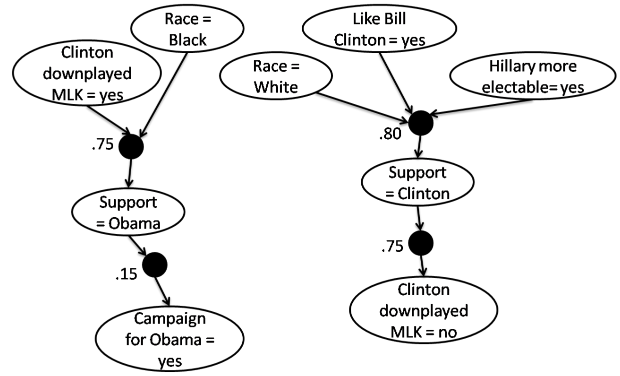

Figure 2: A BKB fragment from a knowledge base about the South Carolina Democratic Primary election as a directed graph.

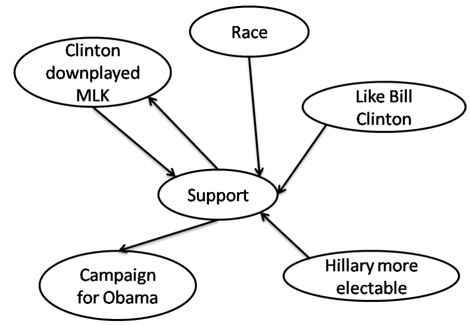

Figure 3: Underlying random variable relationships in Figure 2.

felt that Hillary Clinton had downplayed Martin Luther King Jr.'s role in the civil rights movement and this caused them to support Obama (Shipman 2008). However, some white voters who liked the Clinton family and supported Hillary may have been skeptical that she had any ill intentions when making the remark about Martin Luther King Jr. Similar to d-separation in BNs, it is possible to determine independence semantics from the graph induced by a BKB (Shimony, Santos, and Rosen 2000).

We now give the formal definition of the graphical representation of a BKB from (Santos and Dinh 2008):

Definition 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 $\alpha \in I$ such that $(q, \alpha) \in E$. If there is a link from $q \in S$ to $\alpha \in I$, we say that q *supports* α.

An edge $(a, b) \in E$ will be denoted as $a \to b$ throughout the rest of the paper. For each S-node q in a correlation graph G, we denote the set of all I-nodes that point to q as $Tail_G(q)$, i.e. $Tail_G(q) = \{\alpha \in I | \alpha \to q \in E\}$. Similarly the $Head_G(q)$ is the I-node that q points to in G, i.e. the α such that $q \to \alpha \in E$.

Two sets of I-nodes, I_1 and I_2 are said to be *mutually exclusive* if there is an I-node $(R = v_1)$ in I_1 and an I-node $(R = v_2)$ in I_2 with $v_1 \neq v_2$. Intuitively, mutual exclusivity is the condition that the two sets I_1 and I_2 cannot be satisfiable at one time (i.e. there must be some random variable (in this case R) that is given a contradictory assignment in each of the sets). Similarly, two S-nodes q_1 and q_2 are called *mutually exclusive* if $Tail(q_1)$ and $Tail(q_2)$ are mutually exclusive. S-nodes and I-node sets that are not mutually exclusive are called *compatible*. A *state* is a set of I-nodes θ such that there is at most one instantiation of each random variable in θ.

Definition A *Bayesian knowledge base* (BKB) is a tuple $K = (G, w)$ where $G = (I \cup S, E)$ is a correlation-graph, and $w : S \to [0, 1]$ such that

1. $\forall q \in S$, $Tail_G(q)$ 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, $Tail_G(q_1)$ and $Tail_G(q_2)$ are mutually exclusive
3. For any $Q \subseteq S$ such that (i) $Head_G(q_1)$ and $Head_G(q_2)$ are mutually exclusive, and (ii) $Tail_G(q_1)$ and $Tail_G(q_2)$ are not mutually exclusive for all q_1 and q_2 in Q,

$$\sum_{q \in Q} w(q) \leq 1$$

So a BKB is a correlation graph along with a weight function (specifying the probabilities in the CPRs) satisfying the above conditions. The first condition states that no rule can have two contradictory assignments in its tail. The second ensures that all S-nodes pointing to the same I-node are mutually exclusive, and the last condition ensures normalization of the probabilities.

Reasoning with BKBs takes on two main forms called belief revision and belief updating, similar to reasoning with BNs (Pearl 1988). Belief revision has also been called the maximum a posteriori (MAP) or most probable explanation (MPE). In belief revision we have some evidence in the form of assignments of random variables to states, e.g. assume we have a BKB with n variables and k pieces of evidence $A_1 = a_1, A_2 = a_2, \ldots, A_k = a_k$. The goal of belief revision is to find the assignment of $A_{k+1}, A_{k+2}, \ldots, A_n$ to states such that $P(A_{k+1} = a_{k+1}, \ldots, A_n = A_n | A_1 = a_1, \ldots, A_n = a_k)$ is maximized. Intuitively, this can be thought of as finding the most probable state of the world that contains all the evidence. In belief updating, the goal is to find the probability of a state of a random variable given the evidence. For example, assuming again that $A_1 = a_1, A_2 = a_2, \ldots, A_n = a_k$, we may want to find $P(B = b | A_1 = a_1, A_2 = a_2, \ldots, A_n = a_k)$ for some instantiation b of a random variable B.

In BKBs we are often interested in partial states of the world instead of complete states. In revision, we can compute the most likely partial state of the world (also called the most probable inference as will be described later), and in updating, we can find the marginal probability of a partial state of the world in addition to a single single state of a random variable or a complete state of the world. Given that BKBs support incomplete specification of knowledge, we may have to make these calculations in the face of this incompleteness. In these situations, only the information that has been specified by the knowledge engineer is used in the calculations.

In carrying out both types of reasononing, we make use of subgraphs called *inferences* that capture the joint probability of the I-nodes in the subgraph. Intuitively, an inference is a partial state of the world.

Definition 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*) $\forall \alpha \in I'$, $\exists q \in S'$, $q \to \alpha \in E'$
3. (*Well-founded*) $\forall q \in S'$, $Tail_\tau(q) = Tail_G(q)$
4. (*Well-defined*) $\forall q \in S'$, $Head_\tau(q) = Head_G(q)$
5. There is at most one I-node corresponding to any given random variable in I'.

The properties above state that in an inference, (2) each I-node must have an S-node pointing to it, (3) if an S-node in the inference has any I-nodes pointing to it in the correlation graph they must also be in the inference, (4) there cannot be S-nodes in the inference that do not point to any I-nodes in the inference, and (5) there cannot be mutually exclusive I-nodes in an inference. The joint probability of the set of I-nodes in an inference is simply the product of the weights of all S-nodes in the inference, i.e. for an inference $\tau = (I' \cup S', E')$, $P(\tau) = \prod_{q \in S'} w(q)$.

The Fusion Algorithm

In this section, we will define an algorithm that takes multiple Bayesian knowledge fragments as input and fuses them into one larger BKF. Each BKF is assumed to come from a distinct expert source. The reliability and probability of correctness of each source is assessed to produce a *reliability index* for each fragment allowing more reliable sources to be given more weight in the fused BKF. As long as the input fragments are valid BKBs, the algorithm assures that the fusion of these BKBs is still a valid BKB. A BKF can thus be described as a BKB along with a source and its reliability index, i.e. it can be described as a tuple $K = (G, w, \sigma, r(\sigma))$ where G and w define a valid BKB, σ is the source of the BKB, and $r(\sigma)$ is the reliability index, a non-negative real number denoting the reliability of the source.

The idea behind the algorithm is to first take the union of all input fragments. Then for each random variable A in the fused BKF we create a new random variable S_A representing the source of a rule supporting an I-node representing a state of A. Each S-Node that points to a state of A has an I-node added to its tail called a *source node* whose random variable is S_A and whose state is σ, the source of the fragment the S-node came from. The reliability of this source node is $r(\sigma)$, the reliability of the fragment that the S-node came from. So if a rule q supporting $(A = a)$ is from source σ, we would add the I-node $(S_A = \sigma)$ to the tail of q in the fused BKF. We also add an S-node q' to the fused BKF whose head is $(S_A = \sigma)$. The weight of this S-node is proportional to the reliability index $r(\sigma)$ for the source, normalized so that the weights of all sources for a given random variable cannot exceed 1.

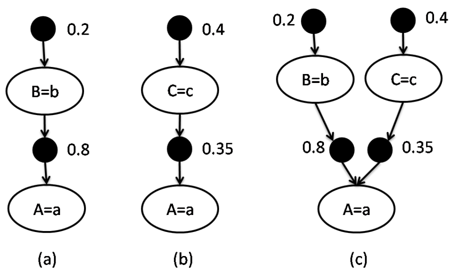

Figure 4: Naive union of two fragments results in a violation of mutual exclusivity.

In addition to allowing us to keep track of sources in the model, the source nodes also resolve problems that could occur if we decided to just take the simple union of several fragments. Consider the two fragments in Figure 4 (a) and (b) whose naive union is shown in 4 (c). We can see that this is not a valid BKB because the two S-nodes supporting $(A = a)$ are not mutually exclusive. If we add in source nodes as in Figure 5 we can see that the S-nodes are now mutually exclusive because the source nodes in the tails of the random variables cannot both be true at the same time. The algorithm also solves normalization problems that may occur after fusion. There is a danger of violating property (3) in the definition of a BKB above and ending up with probabilities that can sum to more than 1 using a naive fusion approach, but the source nodes prevent this. Intuitively, adding a source node turns a rule that says "If A is true then B with probability p" into a rule that says "If source σ is correct and A is true, then B with probability p".

The algorithm can be described more formally as follows. The input is a set of n BKFs, $\{K_1, K_2, \ldots, K_n\}$ where $K_i = (G_i, w_i, \sigma_i, r(\sigma_i))$. The output is a new BKB, $K' = (G', w')$ with $G' = (I' \cup S', E')$, that is the fusion of the n input BKFs. For an I-node α in some fragment, let R_α be the random variable that α is an instantiation of.

BAYESIAN-KNOWLEDGE-FUSION$(K_1, K_2, \ldots, K_n)$

Let $G' = (I', S', E')$ be an empty correlation graph
for all fragments K_i with $i \leftarrow 1$ **to** n
  for all S-nodes $q \in S_i$
    Let $\alpha \leftarrow Head_{G_i}(q)$
    Let the source I-node for q
      be $s = (S_{R_\alpha} = \sigma_i)$
    Add q, all nodes connected to q in G_i,
      and the corresponding edges to G'
    Add s to G' along with an S-node
      supporting it
Let ρ be a normalizing constant
for all S-nodes q' supporting some source node s
  Let $w'(q') \leftarrow r(s)/\rho$
return $K' = (G', w')$

To compute the normalizing constant ρ we compute for each random variable R, the sum of the reliabilities of all source nodes supporting an instantiation of R. ρ is then set to the maximum of these sums.

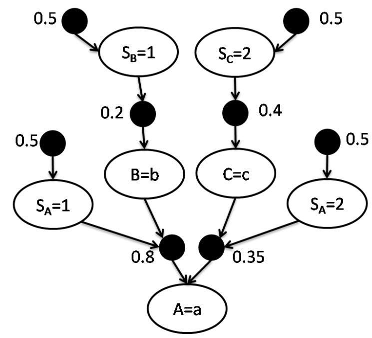

Figure 5: Fusion of the two fragments from Figure 4 using our Bayesian knowledge fusion algorithm solves mutual exclusivity problem.

Algorithm Properties

As long as the fragments that are input into the algorithm are themselves valid BKBs, the output of the algorithm will also be a valid BKB. To prove this we need to show that the conditions in the definition of a BKB are satisfied.

Theorem 1 *The output* $K' = (G', w')$ *of* BAYESIAN-KNOWLEDGE-FUSION *is a valid BKB.*

We now sketch the proof of this theorem:

Proof The fused BKB is clearly a correlation-graph with a weight function, so the other three properies are all that remain to be shown.

1. There is at most one instantiation of each random variable in the tail of any S-node because we added only one I-node to each tail and it is of a random variable that did not previously exist.
2. *(Mutual exclusivity)* Let q_i be any S-node in G' and assume it came from input fragment K_i and has head $(R = v)$. If there are any other S-nodes from K_i with the same head, they are mutually exclusive with q since K_i is itself a BKB and S-nodes with the same head in a BKB must be mutually exclusive. Other S-nodes with the same head that originate from a fragment other than i are still mutually exclusive with q_i because each has a source random variable, S_R in its tail with different states (corresponding to the fragment the S-node originated from).
3. *(Normalization)*: Now we need to show for any set of mutually exclusive S-nodes with compatible tails, the sum of the weights is less than 1. Any such set will have have all S-nodes pointing to I-nodes of the same variable (or else wouldn't be mutually exclusive). Assume that random variable is a source variable, then the normalization step in the algorithm forces the sum of their weights to be less than or equal to 1. If the set points to some other random variable, all S-nodes in the set must be from the same fragment because S-nodes from different fragments would have source nodes in their tail with different instantiations, making the tails mutually exclusive (and so not compatible).

Therefore the result of the fusion algorithm is a valid BKB. ∎

Now that we know the result of the fusion algorithm is a valid BKB if all input fragments are valid BKBs, it is interesting to ask what other properties of the input fragments are preserved by the algorithm. One property that we know is preserved is *groundedness*.

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

Groundedness is a desirable property because if an S-node remains grounded during changes to the knowledge base, the initial probability assigned to it is preserved (Santos, Santos, and Shimony 2003).

Theorem 2 *If all S-nodes in the input fragments are grounded in their respective fragments, then all S-nodes in the fusion of these fragments, K', are grounded in K'.*

We now sketch a proof of this theorem:

Proof Let $q \in S'$ be an S-node originally from fragment K_i, then there is a subgraph τ that is an inference over K_i such that $q \in \tau$. We will construct a new subgraph $\tau' \supset \tau$ that is an inference over K'. For each S-node $q' \in \tau$, the fusion algorithm added a corresponding source node, call it α' to the fused BKB. Add α' along with the edge $\alpha' \to q'$ to the new inference τ' along with α''s source node.

We claim that this new subgraph, τ' is an inference that contains q. It is identical to τ except for the addition of the source nodes and their supports. As a result, it is acyclic as required by the definition of an inference (since τ was acyclic and the source nodes' supports have an empty tail). We only added one I-node and we also added a support for it (so τ' is well-supported). We only added one S-node with no ancestors in the BKB, so τ' is well-founded, but it has one child (the source node), so it is well-defined. Finally all I-nodes added to τ' are from distinct random variables that are not found in τ, so there is at most one I-node corresponding to a given random variable in the BKB. Therefore τ' is an inference. ∎

In addition to groundedness of S-nodes we are also interested in whether I-nodes or sets of I-nodes participate in any inferences.

Corollary 3 *If all S-nodes in the input BKBs are grounded in their respective BKBs, then all I-nodes in K' are grounded in K'.*

Proof Assume α is an I-node in K', then it must be in either the head or the tail of some S-node q. By the previous result, q is in some inference τ' in the fused BKB. Since τ' is well-founded and well-defined, the head and tail of q in K' must also be in τ'. Therefore α is in τ' and so α is grounded. ∎

This establishes that single I-nodes remain grounded in the fused BKB, but what about sets of I-nodes?

Definition A state θ is *well-represented* in a BKB K if there exists an inference $\tau = (I' \cup S', E)$ over K whose I-node set I' coincides with θ, i.e. if $\theta \subseteq I'$.

Similar to groundedness, we can show that the fusion algorithm also preserves well-representedness:

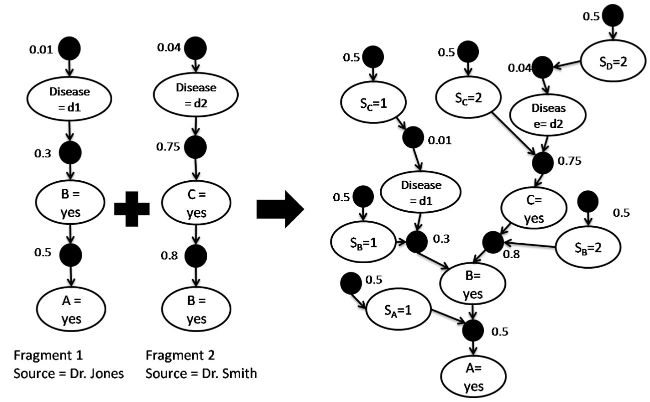

Figure 6: Fragments from two different doctors on the left. The BKB resulting from their fusion is shown on the right.

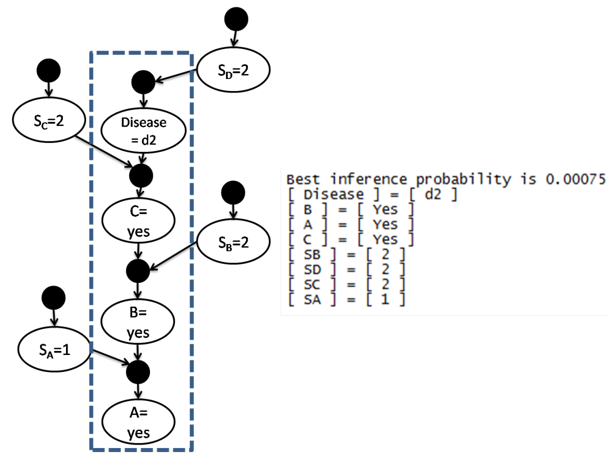

Figure 7: Most probable inference from BKB in Figure 6 with evidence $(A = yes)$ contains the original nodes (center) and the source nodes that were used to support them.

Corollary 4 *All well-represented states in the input fragments are well-represented in K'.*

Proof If a state θ was well-represented in some fragment K_i then it was in an inference τ in K_i. After the fusion process it is also in the inference τ' discussed in proof of the previous theorem. ∎

Examples

Now, an example of how the fusion process can be used to aggregate information from multiple experts will be given in this section. Assume there are two doctors asked to provide rules to a diagnosis system. They do not have much time to help so they each provide only a few simple rules. The provided fragments are shown on the left of Figure 6. The fusion of these two fragments is shown on the right of the same figure. In the fragments, random variables A, B, and C correspond to symptoms, and *disease* is a random variable representing the possible diseases a patient may have.

If we are presented with a patient who has symptom A, we can perform belief revision on the fragments to infer a likely disease. From fragment 1 we would conclude the patient had d_1. From fragment 2 we would not be able to make any conclusions, but from the fusion of the two fragments we would see that d_2 was most likely. This would be done by performing belief revision on each BKB to find the most probable inference containing the evidence that $(A = yes)$. The result of belief revision on the fused BKB in Figure 6 is shown in Figure 7. The setting of each random variable in the most probable inference is shown. Each source random variable is also assigned to a state indicating which source was used to support the truth of the random variable. The output can be interpreted: The most likely symptoms are A, B, and C. The source of the rule supporting A was 1, and the source of the rule supporting B, C, and the disease was 2 in this inference.

The authors have used the fusion process in modeling multiple real world scenarios. In one instance the conflict between gangs and the government in Haiti was modeled. In another the South Carolina Democratic primary election was studied. In both cases many sources were used such as news reports, blogs, and expert analysis. A fragment was created from each of the sources and then they were fused for analysis. The fusion algorithm proved especially helpful in two regards. First it was simply too difficult to work with a network the size of the final fused BKB, it was much easier to deal with each of the fragments separately. Second, there were often multiple violations of the mutual exclusivity condition between any two fragments neccesitating either a change in the fragments or an algorithm such as ours in order to aggregate them.

As a note, the fusion algorithm will work on Bayesian networks as well as BKBs, but when you fuse BNs, you will most likely end up with a BKB, not a BN, after the fusion. Cycles may be introduced violating the acyclicity of BNs, and rules will almost certainly be missing from the fused network making it an invalid BN, but still a valid BKB. However, the resulting BKB will be probabilistically complete in that there will exist a set of mutually exclusive inferences whose probabilities sum to one.

Conclusion and Future Work

In this paper we introduced an algorithm to fuse several Bayesian knowledge fragments into one BKB. This allows the easy aggregation and de-aggregation of chunks of information from multiple sources. The fusion was shown to produce a valid BKB that preserves the rules in the input fragments and their groundedness. The fusion process also allows for the tracking of sources and construction of explanations that contain rules from multiple experts, forming an explanation with greater likelihood than one formed from rules taken from a single expert in isolation. The algorithm allows for disagreement and even circularity in the rules provided by different sources.

Future research in this area will further investigate the properties of fused knowledge bases and also identify possible applications to demonstrate their usefulness. One particular problem that will be addressed is the validation of fused BKBs. Previous work has looked at how to validate the correctness of a single BKB with regard to a set of test cases (Santos 2001; Santos and Dinh 2008). Given a set of fragments that all pass their own individual test cases, we will investigate under what conditions it can be guaranteed that the fused BKB will also be able to pass these test cases. In particular, if the fused BKB does not pass all test cases, is there a way of adjusting the reliabilities of the input fragments so that it does pass?

Acknowledgements

This work was supported in part by Defense Threat Reduction Agency Grant No. HDTRA1-07-1-0003, Air Force Office of Scientific Research Grant Nos. FA9550-07-1-0050 and FA9550- 06-1-0169, and Office of Naval Research Grant No. N00014-06-C-0020. Special thanks to the Air Force Research Laboratory Information Directorate and Securboration, Inc. for their continued support and transition.

References

Koller, D., and Pfeffer, A. 1997. Object-oriented Bayesian networks. 302–313. Morgan Kaufmann.

Laskey, K. B. 2008. MEBN: A language for first-order Bayesian knowledge bases. *Artificial Intelligence* 172(2-3):140 – 178.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Mateo, CA: Morgan Kaufmann.

Rosen, T.; Shimony, S. E.; and Santos, Jr., E. 2004. Reasoning with BKBs algorithms and complexity. *Annals of Mathematics and Artificial Intelligence* 40(3-4):403–425.

Santos, Jr., E., and Dinh, H. T. 2008. On automatic knowledge validation for Bayesian knowledge bases. *Data Knowl. Eng.* 64(1):218–241.

Santos, Jr., E., and Santos, E. S. 1999. A framework for building knowledge-bases under uncertainty. *Journal of Experimental and Theoretical Artificial Intelligence* 11(2):265–286.

Santos, Jr., E.; Santos, E. S.; and Shimony, S. E. 2003. Implicitly preserving semantics during incremental knowledge base acquisition under uncertainty. *International Journal of Approximate Reasoning* 33(1):71–94.

Santos, Jr., E. 2001. Verification and validation of Bayesian knowledge-bases. *Data Knowl. Eng.* 37(3):307–329.

Shimony, S. E.; Santos, Jr., E.; and Rosen, T. 2000. Independence semantics for BKBs. In *Proceedings of the Thirteenth International Florida Artificial Intelligence Research Society Conference*, 308–312. AAAI Press.

Shipman, T. 2008. Obama supporter warns Clinton over 'racism'. *The Daily Telegraph* (January 19).

Sutton, C.; Morrison, C.; Cohen, P. R.; Moody, J.; and Adibi, J. 2004. A Bayesian blackboard for information fusion. In *In Proceedings of the 7th International Conference on Information Fusion.*