Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

On a Framework for the Prediction and Explanation of Changing Opinions

Eunice E. Santos*, Eugene Santos Jr.†, John T. Wilkinson†, Huadong Xia*
*Department of Computer Science
Virginia Polytechnic Institute and State University, Blacksburg, VA 24060
Email: santos@cs.vt.edu, xhd@vt.edu
†Thayer School of Engineering
Dartmouth College, Hanover, NH 03755
Email: {Eugene.Santos.Jr, John.T.Wilkinson}@dartmouth.edu

***Abstract*—One of the greatest challenges in accurately modeling a human system is the integration of dynamic, fine-grained information in a meaningful way. A model must allow for reasoning in the face of uncertain and incomplete information and be able to provide an easy to understand explanation of why the system is behaving as it is. To date, work in multi-agent systems has failed to come close to capturing these critical elements. Much of the problem is due the fact that most theories about the behavior of such a system are not computational in nature, they come from the social sciences. It is very difficult to successfully get from these qualitative social theories to meaningful computational models of the same phenomena.**

We focus on analysis of human populations where discerning the opinions of the members of the populace is integral in understanding behavior on an individual and group level. Our approach allows the easy aggregation and de-aggregation of information from multiple sources and in multiple data types into a unified model. We also present an algorithm that can be used to automatically detect the variables in the model that are causing changes in opinion over time. This gives our model the capability to explain why swings in opinion may be experienced in a principled, computational manner. An example is given based on the 2008 South Carolina Democratic Primary election. We show that our model is able to provide both predictions of how the population may vote and why they are voting this way. Our results compare favorably with the election results and our explanation of the changing trends compares favorably with the explanations given by experts.

***Index Terms*—Bayesian Knowledge Bases, multi-agent system, probability theory.**

I. Introduction

There are many challenges to be overcome in order to successfully model complex social phenomena such as how people decide who they will vote for in an election. There are a myriad of factors at play and a sea of information that must be processed. Traditionally this type of system has been studied in the social sciences using qualitative theoretical models, expert analysis, or techniques such as polling. These have proved insufficient in providing a detailed understanding of such phenomena and we argue that computational approaches are the required solution. Factors such as dynamism and information that is incomplete, imprecise, and uncertain must be captured in any computational model that has a chance to be successful. Computational social science has been gaining ground in recent years, but there are still no modeling frameworks that are able to incorporate all these factors while representing knowledge at a level this is fine grained enough to draw meaningful conclusions and produce informative explanations of the underlying processes, beliefs, and ideologies driving social phenomena.

In this paper, a comprehensive framework is described that provides a representation of a population of humans and the knowledge they possess in the form of a multi-agent system. Their knowledge and the changes in it over time are used to determine their likely opinions, identify the factors contributing most to these opinions, and identify the factors leading to significant changes in opinion. The model is dynamic in that knowledge is added, removed, and modified over time, and algorithms are developed to analyze changes over time. A probabilistic representation known as a Bayesian Knowledge Base (BKB) [1] is used to capture an agent's knowledge. BKBs naturally capture the uncertainty and incompleteness of information in a complex adaptive system such as a human population.

In our framework, a population is generated and a knowledge base constructed for each agent. The agents have varying characteristics such as age, race, religion, etc., and the knowledge that is assigned to them is dependant upon these characteristics. This allows each individual to have their own opinion based on their cultural background and ideology. The framework was applied to a study of the 2008 South Carolina Democratic Primary election. In this case, census data was used to generate the population and knowledge was pulled from sources such as news stories and expert analysis to fill the agents' knowledge bases.

Related computational work has been produced in several different fields. Social network analysts have tried to model these types of phenomena, but the problem is a lack of of a sufficient knowledge representation for the agents in the network or a way to usefully integrate a knowledge representation into the network analysis. Axelrod used cellular automata in simulations of culture and knowledge spread, but again the knowledge representation is far too simple [2]. Santos et. al [3] have used Bayesian Knowledge Bases to model high level agent knowledge, but lacked the ability to easily aggregate and de-aggregate this knowledge over time. Silverman [4] has

978-1-4244-2794-9/09/$25.00

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

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

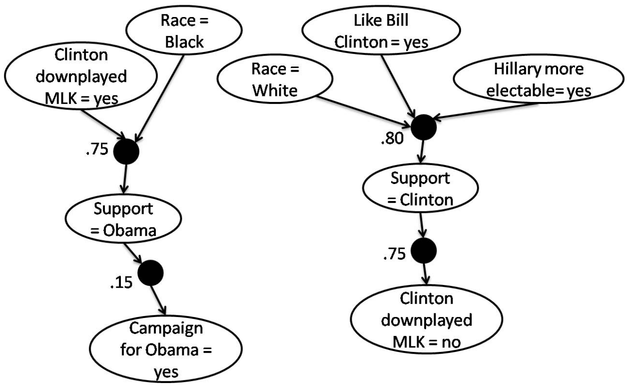

Fig. 2. A BKB fragment from a knowledge base about the South Carolina Democratic Primary election as a directed graph.

developed sociocultural simulations attempting to incorporate sound social theories, but his techniques lack the ability to represent uncertainty and complex interactions in knowledge that our probabilistic framework allows.

In section 2, Bayesian Knowledge Bases will be described along with the methodology used to aggregate them. Then our modeling framework will be presented in detail, followed by a description of the application of this framework to the South Carolina primary. The results of this application will be discussed in the next section followed by a conclusion and future directions for this research.

II. Bayesian Knowledge Bases and Knowledge Fusion

A. 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).

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

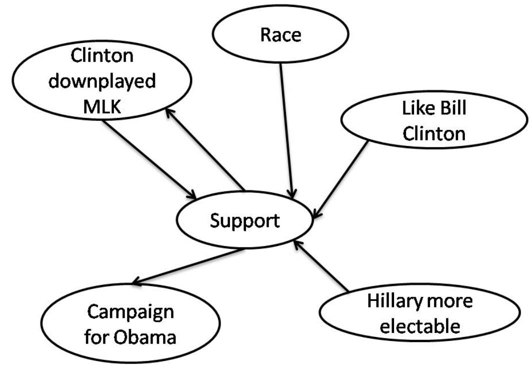

Fig. 3. Underlying random variable relationships in Figure 2.

a BKB built from a news story [5] about the South Carolina Democratic Primary election is shown in rule form in Figure 1 and in graph form in Figure 2 [6]. 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 exists because black voters felt that Hillary Clinton had downplayed Martin Luther King Jr.'s role in the civil rights movement and this caused them to support Obama. 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 [7].

We now give the formal definition of the graphical representation of a BKB from [8]:

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-node sets and I-node sets that are not mutually exclusive are called *compatible*.

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

B. Bayesian Knowledge Fusion

An algorithm that is able to *fuse* a set of BKBs into one larger valid BKB was presented in [6]. In the context of knowledge fusion we refer to each BKB as a *Bayesian knowledge fragment* or simply a *fragment*. Given a set of n fragments $\{K_1, K_2, \ldots, K_n\}$ with $K_i = (G_i, w_i)$ where the source of fragment i is σ_i and the reliability of source σ_i is $r(\sigma_i)$, the fusion algorithm proceeds as follows:

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

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

On line 6, the algorithm uses a source random variable S_{R_α}. One source random variable is created for each I-node α that is the head of an S-node. All S-nodes pointing to α will have a source node in their tail representing some instantiation of the random variable S_{R_α}. The value of $r(\sigma_i)$ is allowed to be any real number. The constant ρ must be chosen such that the normalization condition in the definition of a BKB is satisfied.

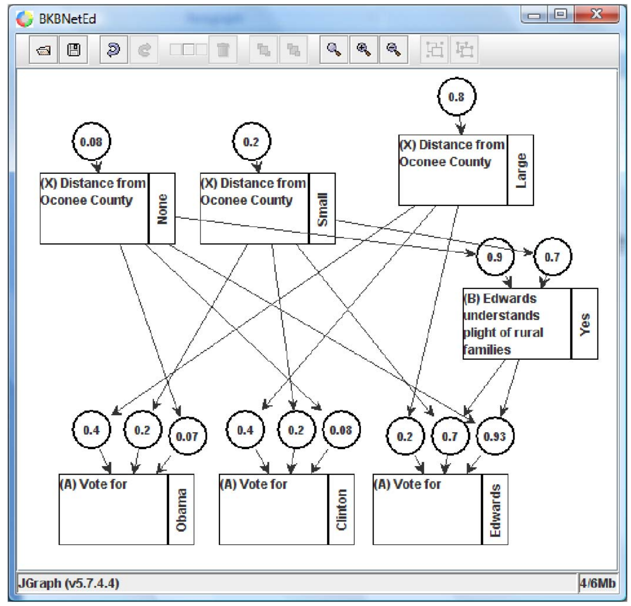

Fig. 4. A fragment describing the effect of a visit by John Edwards to Oconee county.

III. Modeling Framework

The scenarios we wish to model are inherently dynamic. People gain knowledge over time. This knowledge may cause them to modify or even discard previously held beliefs. Thus a key element in our framework is the ability to easily add and remove knowledge over time, as well as to modify the strength of a belief over time. To this end, a global *library of fragments* is maintained which contains all fragments used for all agents in the simulation. Each agent then has their own knowledge base which consists of a subset of the BKFs in the library. Each fragment carries with it information about what type of agent it applies to, similar to metadata. Accordingly, each agent has a set of attributes describing itself. At each time step, a BKF may be added or removed from the knowledge base of the agent based on a matching algorithm which compares fragment metadata to agent attributes in order to choose the appropriate fragments from the library. A simpler version of this process in which time was not taken into consideration was described in [9].

An example of how agent characteristics can affect fragment selection is the case of John Edwards' visit to Oconee county during his campaign. A fragment representing this visit, shown in Figure 4, was only distributed to individuals close enough to the county to be affected by his visit. The weight of this visit in an agent's decision making process lessened the further they were from Oconee by way of the "Distance from Oconee County" random variable. This demonstrates that even if two agents both have the same fragments in their knowledge base, the same reasoning results may not be produced (since random variables in the fragments can be dependent upon agent characteristics).

The weight of the source nodes used in the fusion algorithm represents the *reliability* of the information in each fragment. In our framework, the reliability is then modified based on how long the knowledge has been present in the agent's knowledge base. Recent information is given more weight and the strength of old information fades over time if it is not reinforced. Suppose a fragment f_i enters an agent's knowledge base at time i. The reliability of this fragment at time $j > i$ is computed to be $r(f_i, j) = \frac{1}{j-i+1}$ Thus the reliability of a new fragment is 1 and this fades toward 0 as time goes by.

A. Reasoning in the Framework

Reasoning at a single time step in our model can be accomplished with traditional BKB inference algorithms as discussed in [10]. These include belief revision (also called Maximum A-Posteriori or MAP) and belief updating [11] as in Bayesian networks. What these traditional techniques cannot accomplish however is reasoning about changes that occur between time steps and why these take place, which was one of our major goals in this effort. These techniques can tell us who an agent is likely to vote for but not the factors most influential in determining this agent's vote.

B. Calculating Opinions

The opinion of each agent was calculated using belief updating at each time step. In belief updating we are given some evidence $A_1 = a_1, A_2 = a_2, \ldots, A_k = a_k$ where A_i is a random variable and a_i is one of its states. The goal is to compute $\Pr(B = b | A_1 = a_1, A_2 = a_2, \ldots, A_k = a_k)$ for some random variable B and one of its states b. In order to capture aggregate opinions, and since the results of belief updating are probabilistic, Monte Carlo simulations are used to generate a distribution over group opinions.

C. Contributing Factors

As was mentioned previously, we don't just want to calculate opinions, we want to calculate the factors most influential to these opinions. In order to do this, we applied an algorithm that, given a target I-node q and some other I-node r, computes the *contribution* of node r to the probability of q. We note that the probability of q can be found by looking at all possible "worlds" sanctioned by the BKB and summing the probabilities of all worlds in which q is true. Here, a world is simply an assignment of all variables to a state. Let the set of all the worlds in which q is true be W. Then, in order to find the contribution of r to q, we look at the subset $W' \subseteq W$ in which r is also true. The contribution of r to q is the sum of the probabilities of the worlds in W'.

D. Explaining Changes of Mind

One of the main goals of this work is to be able to provide a computational framework in which to investigate how and under what circumstances people may change their mind. Thus dynamism must play a key role in the analysis. Over time fragments enter and leave an agent's knowledge base and may undergo changes in relevance. We view each new fragment that enters as having some chance of changing the opinion of the agent. This may be through introducing brand new knowledge, linking ideas that the agent previously had not connected to one another, or causing the agent to rethink previously held beliefs.

In the previous section we showed how to compute the contribution of one node to the truth of another in one time step of the model, but contributions can also be used to explain changes in results over time in addition to explaining static results. The idea is that if a sudden change is observed in the system, the algorithm for computing contributions is used to analyze the contributing factors that may have lead to the change.

To accomplish this, a variable in the system is chosen and its contributing factors are measured over time. Specifically, if we desire to compute the factors contributing to changes in an agent's opinion with respect to an I-node q between times $t-1$ and t, we let B_i be the agent's knowledge base at time i. Then the following algorithm is applied.

CONTRIBUTING-FACTORS(t, q)
for all I-nodes $r \neq q$ in B_t
    Let the contribution of r to q at time
        t be $c_r^t(q)$.
    The change in contribution is computed as
    $\delta_q^t(r) = abs(c_q^t(r) - c_q^{t-1}(r))$
Normalize the values of δ so they have a maximum
of 1
Choose a threshold α
    for all values of $\delta_q^t(r)$
    if $\delta_q^t(r) > \alpha$
        then state r is a significant contributor
            to change in q at time t.

If an I-node r is present in the model at time t, but not at time $t-1$, then $c_q^{t-1}(r)$ is set to 0. Similarly, $c_q^t(r) = 0$ if it is present at $t-1$ but not at t. The choice of α here will depend on how significant the system builder would like changes to be in order for them to be reported. We have used trial and error to determine a reasonable value for α thus far. The directionality of the change in contribution is also significant here. An I-node can be considered a significant contributor to a change both by raising its level of contribution over time or by lowering it. A lowering of contribution is significant because it indicates that a variable that previously played a major role in determining an agent's opinion with respect to another variable is no longer sufficient to support the truth of the latter variable.

IV. South Carolina Primary Election

A. Setup

In order to test the framework, a scenario was needed that provided a large amount of easily accessible data. The 2008 South Carolina Democratic Primary was chosen because it met this criteria and there were significant swings in opinion throughout the race. The main candidates were Hillary Clinton, Barack Obama, and to a lesser extent, John Edwards. Other candidates did not receive a significant portion of the vote in

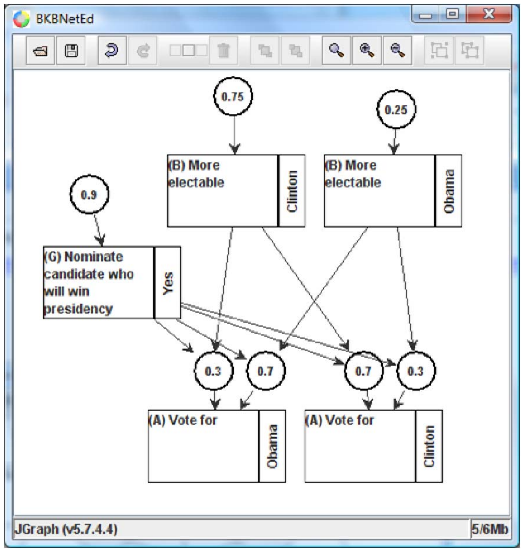

Fig. 5. Example of a fragment from the library showing that Clinton was considered more electable.

polling data or the general election, so we focused our analysis on only the three candidates mentioned above. In the early stages of the election, Hillary Clinton had a large lead in the polls. Through the course of the race, Obama steadily caught up and then went on to take a large lead. The goal of this experiment was to explain this swing.

In constructing the model the major components that needed to be built were the population of agents and the library of fragments. In order to build the population, US census data was used. The data gave demographic information such as age, race, gender, and town of residence. The population of South Carolina is over 4 million, but a smaller representative sample was generated with each town having 0.05% of its true population for a total of around 2000 agents.

The library of fragments was populated with a total of 25 fragments spanning in time from September 27, 2007 to January 22, 2008. An example of one of the fragments is shown in Figure 5. It is from early on in the race and describes the fact that Hillary Clinton was considered more "electable" than Obama. In other words, it was believed that she had a better chance of winning the general election than Obama. Over time this belief faded however as Obama won primaries in other states and issues such as race started to play a key role in people's opinions, as described in the fragments shown previously in Figure 2 and the one in Figure 6.

The actual primary election took place on January 26, 2008. There were 16 different points in time at which fragments were added or removed and their reliabilities updated according to the formula presented earlier. The fragments came from a combination of news stories, expert analysis articles, and polling data. The amount of polling data included was kept to a minimum since one of our goals was to see how the predictions from our model compared to those produced by

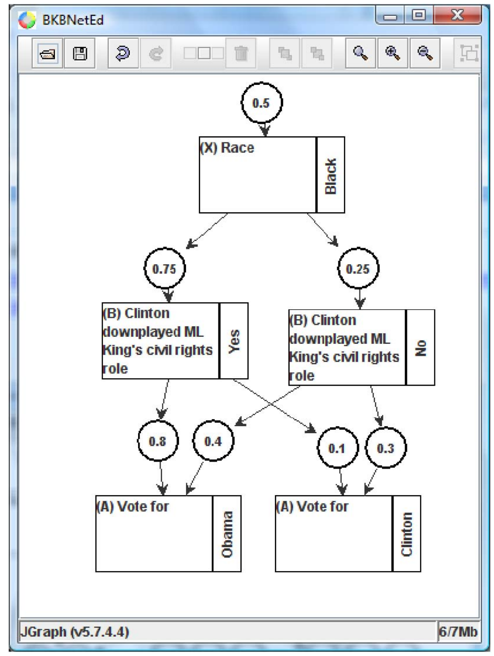

Fig. 6. Example of a fragment from the library describing racial tension in the primary.

the pollsters. Two sets of analyses were run. In one, three of the 16 time points consisted of polling data from the state of South Carolina during the time period mentioned above. In the other, only one set of national polling data from July 7, 2007 was used to seed the model and then all further inputs were purely from news and expert analysis to further minimize the impact of the polls.

Once the population and fragments were constructed, it was possible to apply our algorithms and analyze the results. For each agent, we performed belief updating at each time point. This produced a probability that each agent would vote for each of the three candidates. To then turn this into a prediction of the number of people who would vote for each candidate, we used a Monte Carlo simulation to generate a distribution over the possible votes for each candidate. This provided an idea of the number of people who might vote for the candidate, as well as an idea of the uncertainty in these predictions. The result of the Monte Carlo simulation for Clinton and Obama at one point in time is shown in Figure 7.

Finally we applied our significant contribution algorithm to detect which variables were driving the changes we were seeing in the voters' opinions. This was done separately for each of the three candidates.

B. Results

A website called Real Clear Politics[1] provides an up to date average of polling data for a variety of political races. We took Real Clear Politics' averages on the 16 time points included in

[1] http://www.realclearpolitics.com/

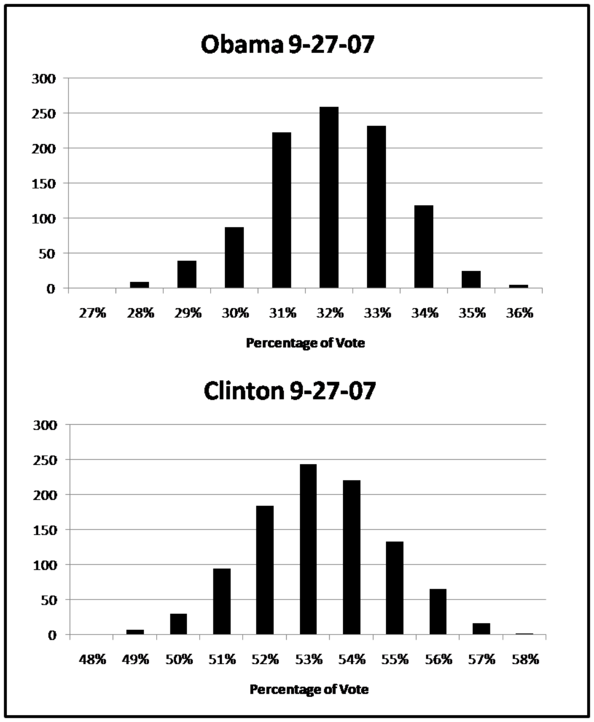

Fig. 7. Monte Carlo simulation results for Clinton and Obama for one date.

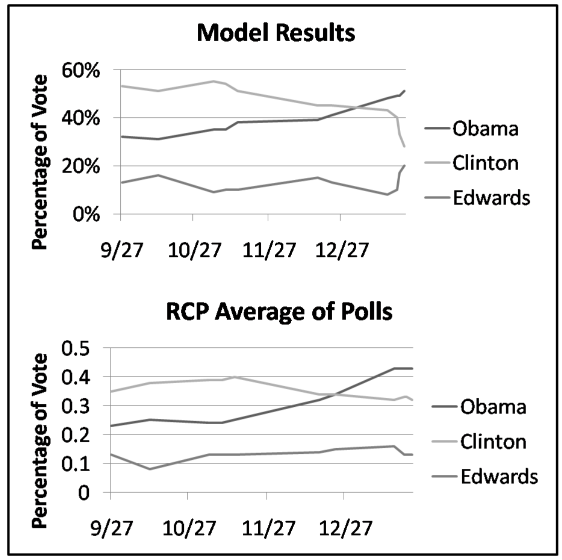

Fig. 8. Model results compared to polling results.

our model and compared them to the predictions generated by our model. In the scenario where three time points contained polling data from South Carolina, we observed that the trends seen in the polling data were very similar to those produced by the model, even in between the injection of new polling data. Results from the model and Real Clear Politics are shown in Figure 8.

Furthermore, the final prediction from our model (on January 22, four days before the election) was closer to the final election results than the average of the polls on the same day. In the scenario where only a national poll from seven months before the election was used to seed the model, we still observed very similar trends to the polling data although the final prediction was slightly further away from the election results than the final polling average.

The significant contribution algorithm was more difficult to validate as there is no clear ground truth as to which variables had the biggest impact on the election. In order to see that the results produced were at least reasonable, we searched for political pundits who had commented on the race after its conclusion to see if they cited similar variables as significant factors to those identified by our model. In general we were able to find pundits who agreed with most of the factors our algorithm highlighted. This is a good first step toward validation of the algorithm.

Some of the results from the algorithm that had a high significance are shown in Figure 9. One of the significant issues identified by the algorithm was whether Obama or Clinton was more electable. Thus the model is indicating that changes in people's beliefs regarding this issue played a major role in the actors' voting decisions. This fact was also cited by several pundits after the election as a major factor. For instance, one analyst said "South Carolina's black voters seemed to doubt Obama's ability to win the presidency. Today, those voters propelled him into a solid victory over Clinton in this Democratic primary."[2].

Another significant issue our algorithm identified was the racial tension that ensued after Clinton was perceived to have slighted Martin Luther King Jr. This was also cited by several sources after the election as a major factor. For instance, the New York Times reported "supporters of Mrs. Clintons campaign and chiefly, her husband were accused of racially tinged attacks and innuendo against Mr. Obama before the South Carolina primary. Mr. Obama went on to rout Mrs. Clinton on the strength of strong support from blacks, a constituency Mrs. Clinton had courted hard."[3]

C. Discussion

It is very difficult to validate a model such as this one. One concern is that we use polling data as a part of our input (in the form of a BKB) and then compare the output to polling data. As a result we tried to keep the use of polling data to a minimum, but even then, the authors of the news articles and the expert analyses we used most likely had an idea of what the polling numbers were when they wrote their pieces. Thus comparison against polling data cannot serve as the only means of validating such a model. Additionally we imagine that in many of the scenarios that this type of analysis would

[2]http://thegate.nationaljournal.com/2008/01/obama_pulls_off_decisive_win_i_1.php

[3]http://www.nytimes.com/2008/02/02/us/politics/02race.html?em&ex=1202101200&en=76022afec454ec29&ei=5087%0A

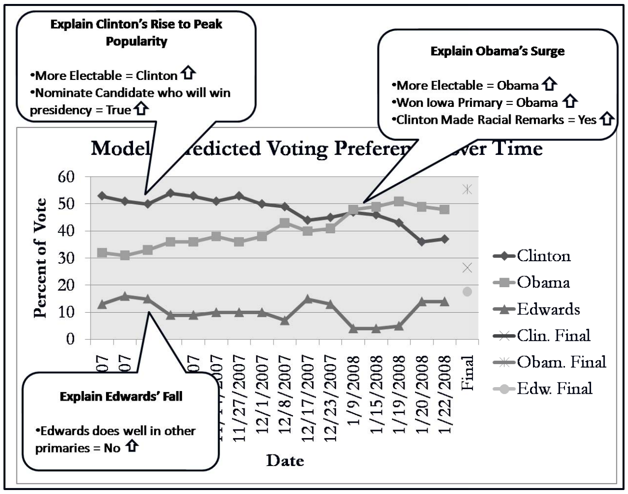

Fig. 9. Some of the contributions identified as significant.

be most useful in, such as conflict situations where the opinions of the actors involved may have a substantial impact on the outcome, polling data will not be available for validation or input purposes.

For this reason, one of our main focuses in this work is not only to provide predictions, but also to *explain* why those particular predictions are produced. Unfortunately this also poses a great challenge in terms of validation. It's not clear how a mathematically rigorous evaluation of this capability could be conducted. Despite these concerns, we believe that the results presented do show that our framework has great promise in helping to analyze, understand, and anticipate the opinions of the agents populating the system.

V. Conclusion

A multi-agent system based technique was presented to aid in analyzing the likely opinions of the agents in the system and to calculate the factors exerting the most influence over these opinions. An algorithm was also detailed allowing for the detection of variables causing significant changes in opinion over time. Experimental results based on the 2008 South Carolina Democratic Primary election show that the system has promise in the modeling of real world scenarios, however there is much work still to be done.

Currently the agents exist and think in isolation of one another. An obvious extension to this work is to consider the effects of an agents opinion on their friends, family, and acquaintances. In other words to study the effects of an agent's *social network* on their beliefs. Going a step further, the model needs to be able to capture how information moves throughout the social network. Diffusion models have been studied in fields such as disease spread, and some have also attempted diffusion models in social networks. Unfortunately these models lack a sufficiently powerful representation of the knowledge the agents in the social network possess. Without considering an actor's prior knowledge, beliefs, and goals it is impossible to come up with an accurate model of what knowledge they will accept and integrate into their knowledge base, what impact this new knowledge may have on their decision making processes, and what knowledge they will simply reject.

Acknowledgment

Thanks to Scott K. Wong for his insightful input. This work was supported in part by Defense Threat Reduction Agency Grant No. HDTRA1-07-1-0003 and Air Force Office of Scientific Research Grant Nos. FA9550-07-1-0050 and FA9550-06-1-0169.

References

[1] E. Santos, Jr. and E. S. Santos, "A framework for building knowledge-bases under uncertainty," *Journal of Experimental and Theoretical Artificial Intelligence*, vol. 11, no. 2, pp. 265–286, 1999.

[2] R. Axelrod, "The dissemination of culture: A model with local convergence and global polarization," *The Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997. [Online]. Available: http://dx.doi.org/10.2307/174371

[3] E. J. Santos, Q. Zhao, F. Pratto, A. R. Pearson, B. McQueary, A. Breeden, and L. Krause, "Modeling multiple communities of interest for interactive simulation and gaming: the dynamic adversarial gaming algorithm project," in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 6564, May 2007.

[4] B. G. Silverman, G. Bharathy, M. Johns, R. J. Eidelson, T. E. Smith, and B. Nye, "Sociocultural games for training and analysis." *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, vol. 37, no. 6, pp. 1113–1130, 2007.

[5] T. Shipman, "Obama supporter warns Clinton over 'racism'," *The Daily Telegraph*, no. January 19, 2008.

[6] E. Santos, Jr., J. T. Wilkinson, and E. E. Santos, "Bayesian knowledge fusion," in *Proc. 22nd International FLAIRS Conference*. Sanibel Island, FL: AAAI Press, 2009, pp. 559–564.

[7] S. E. Shimony, E. Santos, Jr., and T. Rosen, "Independence semantics for BKBs," in *Proceedings of the Thirteenth International FLAIRS Conference*. AAAI Press, 2000, pp. 308–312.

[8] E. Santos, Jr. and H. T. Dinh, "On automatic knowledge validation for Bayesian knowledge bases," *Data Knowl. Eng.*, vol. 64, no. 1, pp. 218–241, 2008.

[9] E. E. Santos, E. S. Jr., L. Pan, and J. T. Wilkinson, "Culturally infused social network analysis," in *IC-AI*, 2008, pp. 449–455.

[10] T. Rosen, S. E. Shimony, and E. Santos, Jr., "Reasoning with BKBs algorithms and complexity," *Annals of Mathematics and Artificial Intelligence*, vol. 40, no. 3-4, pp. 403–425, 2004.

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