A Temporal Policy for Trusting Information

Karen K. Fullam and K. Suzanne Barber

Laboratory for Intelligent Processes and Systems,
The University of Texas at Austin,
Austin, TX 78712
{kfullam, barber}@lips.utexas.edu

Abstract. In making a decision, an agent requires information from other agents about the current state of its environment. Unfortunately, the agent can never know the absolute truth about its environment because the information it receives is uncertain. When the environment changes more rapidly than sources provide information, an agent faces the problem of forming its beliefs from information that may be out-of-date. This research reviews several logical policies for evaluating the trustworthiness of information; most importantly, this work introduces a new policy for temporal information trust assessment, basing an agent's trust in information on its recentness. The belief maintenance algorithm described here values information against these policies and evaluates tradeoffs in cases of policy conflicts. The definition of a belief interval provides the agent with flexibility to acknowledge that a belief subject may be changing between belief revision instances. Since the belief interval framework describes the belief probability distribution over time, it allows the agent to decrease its certainty on its beliefs as they age.

Experimental results show the clear advantage of an algorithm that performs certainty depreciation over belief intervals and evaluates source information based in information age. This algorithm derives more accurate beliefs at belief revision and maintains more accurate belief certainty assessments as belief intervals age than an algorithm that is not temporally-sensitive.

1 Introduction

In making a decision, an agent requires knowledge about the current state of its environment. When the agent is part of a multi-agent system, it often must communicate with other agents to obtain information for building the agent's environment model. Unfortunately, the agent can never know the absolute truth about the current state of its environment because the information the agent collects about its environment is uncertain. There is uncertainty in the agent's perceptions and communications with other agents, here called information sources, who may be incompetent or dishonest. However, the agent can model its uncertain environment as a set of beliefs, or consistent estimates the agent asserts to be true. Belief maintenance is the process by which an agent builds its model of the environment, based on the uncertain information it receives from its sensors and through communication with information sources.

R. Falcone et al. (Eds.): Trusting Agents, LNAI 3577, pp. 75 – 94, 2005.

This research examines the trustworthiness evaluation of information among agents in a multi-agent system; we highlight the perspective of a single agent, who receives information from other agents serving as information sources. Numerous researchers ([2], [6], [7], [11]) have focused on trust between agents, maintaining trust or reputation models of information sources. Although this research does make use of information source trust models, we utilize a different approach: assessing trust in information itself. Information that is trusted is allowed by the agent to impact its belief calculation. This research enumerates several policies for evaluating trust in information, highlighting a single policy related to temporal change in the agent's environment.

Temporal dynamics within an agent's environment complicate the information uncertainty problem. When the true state of the agent's environment changes more rapidly than sources provide information about the environment, an agent faces the problem of forming its beliefs from information that may be out-of-date. Additionally, if the agent must wait long periods of time between belief revision instances, it must have some mechanism for modeling how the environment may be changing while it receives no new information. Source unreliability and information age are two of many causes of uncertainty an agent faces when determining what information to trust. The agent must assess tradeoffs between causes of uncertainty, asking questions such as: should recent information from an unreliable source be trusted more than older information from a more reliable source?

Roorda, et al. [9] criticize AGM theory for giving priority only to the most recent incoming information. Instead, they claim that priority should be given to sources in order of estimated trustworthiness. This research reviews several logical policies for prioritizing information based on trust by examining the reliability of the information source and the content of the information. Most importantly, this work introduces a new policy for temporal information trust assessment, basing an agent's trust in information on its recentness. The belief maintenance algorithm described here evaluates incoming information against these policies and assesses tradeoffs in cases of policy conflicts. Information deemed most trustworthy is given priority when information is merged to form beliefs.

Examining how information and beliefs become untrustworthy over time is important for several reasons. First, when forming a belief, the agent is able to assess the trustworthiness of information provided by the source, determining whether information is "recent enough" to impact belief revision. Second, the agent is able to dynamically assess the certainty of its own beliefs for decision-making purposes, since as time passes, beliefs may become obsolete. A more accurate belief certainty assessment can be calculated when the agent models the age of its source information against its model of how fast the environment changes. Accurate certainty assignment is important for decision-making; in cases where a minimum threshold of belief certainty is required to take action, unwise actions can be avoided if certainty is decreased to identify when desirable action conditions expire. Thus, good decisions can still be made with information that is incomplete or not fully up-to-date. Finally, using a temporal policy can improve scalability of belief revision. The agent designer might choose to set minimum, temporally-based relevance thresholds, eliminating the need to further evaluate source information deemed obsolete and no longer relevant.

The remainder of this paper is organized as follows. Section Two defines belief representations required to accommodate the temporal factors discussed in this research, including the notion of a belief interval. In Section Three, information valuation policies are outlined, with discussion devoted to the inclusion of a policy for valuating information according to recentness. A policy-based belief revision algorithm is described in Section Four, while the experimental results validating the algorithm are discussed in Section Five. Section Six concludes by summarizing this research and identifying paths for future work.

2 Belief Representations to Accommodate Temporal Factors

For the purposes of this research, an abbreviated explanation of the belief representation as formally defined in [1] will be given. Beliefs are based on a belief subject v, where v is any variable or statement that can evaluate to some continuous or discrete set of values. At any point in time, x_T represents the true value of v; however, because of uncertainty, an agent can never exactly know x_T. Therefore, the agent's belief about v is its best estimate of x_T. Let X_a^v be called the set of all possible values for v believed by a to exist. Then, at a time instant t, belief B_t about v is uniquely defined by the tuple $< v, X_a^v, \phi_a^v, a, t>$, where ϕ_a^v is a set of probabilities asserted by a over X_a^v. The complexity of the ϕ_a^v representation determines the precision of the belief while weighing computational requirements. For the algorithm presented in this research, we limit our discussion to continuous-variable beliefs, choosing to model ϕ_a^v as a normal distribution described by a mean (μ_v) and standard deviation (σ_v). This simple representation allows easy computation, though precision is sacrificed when data is not normally distributed.

In the belief definition proposed in [1], certainty on a belief is implicitly maintained by the set of probabilities ϕ_a^v. In cases where X_a^v is a continuous set, certainty can be defined as the asserted probability that the true value of v (x_T) is within a specified small radius δ of a given value x'. Formally, for a given $x' \in X_a^v$, certainty that v's true value, x_T, is in the range $x' - \delta$ to $x' + \delta$ is given by:

$$\text{certainty}\left(x',\delta\right) = \int_{x'-\delta}^{x'+\delta} \phi_a^v\left(x\right)dx \tag{1}$$

for a selected interval radius δ.

Traditional belief revision theory has emphasized how incoming information serves as a trigger for an agent to revise its belief; the agent then holds that same belief until new information arrives to trigger another belief revision. However, an agent need not maintain the same belief between belief revisions. In fact, when the belief subject v is known to be dynamic between discrete belief revision instances, the agent should not maintain a constant belief. Instead, the agent might use predictors, based on its model of how v changes over time, to change its belief even when no new information is received. Therefore, we will define concepts related to belief revision, including the notion of a belief interval, which represents temporal dynamics between belief revisions.

For now, we shall define belief revision simply as some process by which an agent a utilizes a set of information source reports C_a^v at time t to derive a belief B_t about v. (A communicated source report c_{s_i} is of the same format as a belief, but may or may not represent the source's true belief about v). The jth belief revision is performed at time t_j, producing a belief B_{t_j} and instantiating a new belief interval I_j. A belief interval I_j is defined by the tuple $I_j = <v, X_a^v, \Phi_a^v, a, t_j>$, where Φ_a^v, the set of ϕ_a^v probability distributions over t such that $t_j \leq t < t_{j+1}$, defines all beliefs in belief interval I_j. Φ_a^v may be an infinite set when time is considered continuous; when time is measured as discrete steps, Φ_a^v is finite. The belief interval I_j is considered valid for $t_j \leq t < t_{j+1}$, that is, until the next belief revision is performed. Tawfik and Neufeld [10] represent time in Bayesian networks by expressing probabilities as functions of time. However, since beliefs are represented as logical statements, probability functions are two-dimensional, plotting only the probability of the most likely value for v over time. Since we consider beliefs to include a probability distribution expressing the likelihood of all possible values for the belief subject, our belief change function is three dimensional. Figure 1 illustrates a three-dimensional probability function Φ_a^v over time, in which Φ_a^v is held constant until belief revision occurs.

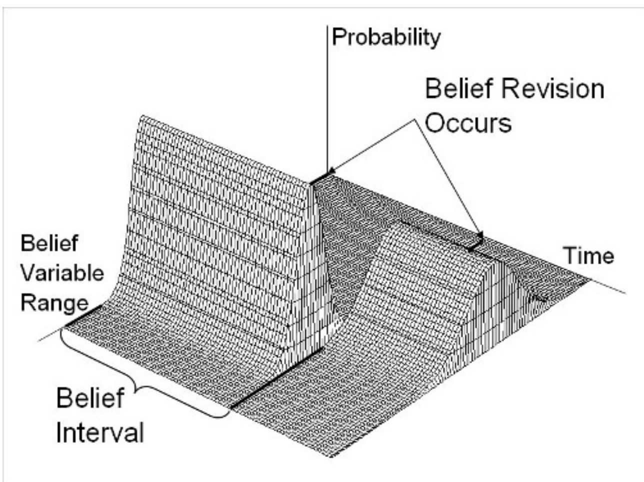

Fig. 1. Belief interval probability distributions over time. The probability distribution function demonstrates discrete changes at belief revision instances

Not only does belief revision generate a new belief B_{t_j}, but it also produces the probability function over the time t for which $t_j \leq t$. At time t_{j+1}, agent a performs belief revision again, generating a new belief interval I_{j+1} from a new set of source reports: $I_{j+1} = <v, X_a^v, \Phi_a^v, a, t_{j+1}>$.

Instant t_{j+1} of belief revision is algorithm-dependent, but may occur after a given number of new source reports are received, or after a given length of time has passed. Regardless of the revision conditions chosen, new information must be received before belief revision can be performed; at least one report in C_a^v must be new. For this research purpose, agents perform belief revision each time a new source report is received. For ease of notation, we shall refer to the belief probability distribution as ϕ_B (described by (μ_B, σ_B)) for the remainder of this paper.

Based on the framework of belief intervals, belief probability distributions can change between belief revisions instants; however, these changes are based on predictive functions. In this paper, we make no attempt to predict changes in the mean of a belief distribution, via trend analysis or any other means. Tawfik and Neufeld [10] build predictive probability distributions, but require domain-specific models of future truth. Similarly, Nodelman, et al. [8] maintain models of expected state changes, but domain knowledge is required to populate those models. Instead, we explore a domain-independent idea of certainty depreciation, in which the standard deviation of the distribution increases over time to reflect increasing uncertainty as beliefs age. We define the age of source information according to the time at which it is initially received; likewise, we define the age of an interval-instantiating belief according to the time at which belief revision occurs.

3 Temporal Information Trust Assessment Policy

Since an agent does not know ground truth, it cannot assess the trustworthiness of information based on a source's communicated values alone. However, the agent can follow general policies for targeting the most accurate information received from sources. These guidelines are naïve heuristics when implemented in isolation; however, an algorithm employing a compromise of all policies can identify robustly the most valuable information for forming beliefs. Each policy assumes source reports are statistically independent. Previous research by the authors [4] has identified several policies for prioritizing trust in information based on:

1. **Priority of Maximum Information:** An agent should incorporate information from as many sources as possible. This concept works best with many reporting sources and requires that sources are statistically independent in their reporting.
2. **Priority to Corroborated Information:** An agent should give priority to information that can be corroborated. Information that is similar to other information should be deemed more trustworthy.
3. **Priority for Source Certainty:** An agent should give priority to information from sources conveying high certainty on that information. If the information source is proficient and honest in conveying a quality certainty assessment, that certainty assessment will be an indication of the trustworthiness of the information.
4. **Priority to Reliable Sources:** An agent should give priority to information from sources it estimates to be most reliable. If a reporting source is estimated by the agent to be a provider of quality information, based on past experience or recommendations from other entities (in other words, the source is considered

reliable), then the agent should deem the information provided by that source more trustworthy.

5. **Priority to Recent Information:** An agent should give priority to information it estimates to be most recent. Since the truth about a belief subject is more likely to have changed as time passes, older information is less likely to be accurate. In order to assess the relative trustworthiness of information with different ages, the agent must know the rate at which truth about a belief subject changes; the faster truth changes, the more quickly information becomes inaccurate.

Fullam and Barber [4] detail belief revision algorithms driven by various subsets of the first four policies, as well as an algorithm which weighs tradeoffs between all four. This research examines the consequences of the fifth policy, which assesses the estimated trustworthiness, or accuracy, of information based on its recentness. The temporal policy introduces a new set of policy tradeoffs which must be addressed. For example, how is older information from a reliable source trusted as compared to more recent information from an unreliable source? When is information too obsolete to be trusted, despite the goal of including maximum information? Is information trustworthy if it is old, even if the source was very certain when the report was initially sent? This research seeks to quantify information trustworthiness in terms of all five policies to find the boundaries between these tradeoffs.

4 Algorithm for Temporal Information Trust Assessment

Previous work assesses trust in information by policies related to either the reliability of the source or by the content of the information [4]. In this research, we incorporate a temporal policy, assessing information trustworthiness based on information age. In addition, we examine the resulting tradeoffs between a temporal policy and the other policies described above. Error due to information age is not the only type of error that must be factored into the agent's trust assessment of source information. Ultimately, we recognize possible information error of three types:

1. **Error due to the age of the information:** As discussed previously, the true value of the belief subject may have changed since the information was received, and we assume the amount of change is related to the amount of time that has passed.
2. **Error communicated by the source:** The source report definition described in this research is flexible, permitting sources to communicate data as a distribution with implicit certainty information. An information source may be uncertain about the information it provides due to the uncertain quality of its own sources or the age of its own information, for example. The magnitude of revision to the source's reliability model is directly related to the certainty conveyed by the source; sources communicating greater certainty experience greater loss or benefit to their revised reliability models. Therefore, sources seeking to maintain the agent's trust have an incentive to accurately communicate their certainty on the information they provide. Because information sources are permitted to convey their own certainty in their reports, the agent is relieved from evaluating the "history" of the information prior to its receipt by the agent itself. For example, the agent need not care about

the age of the data from the source's perspective or the quality of the source's information providers; those factors should be expressed in the source's certainty conveyed to the agent.

3. **Error due to source unreliability:** Information sources may be malicious or incompetent; the degree of estimated source trustworthiness is represented in an agent's source reliability models.

Temporally-caused error must be merged with estimations of information inaccuracy due to both the uncertainty communicated by the source and the unreliability of the source itself. While error due to source uncertainty is easily obtained from the source report, modeling error due to information age and error due to source unreliability require significantly more effort. Therefore, these two types of error are discussed in more detail in the following sections.

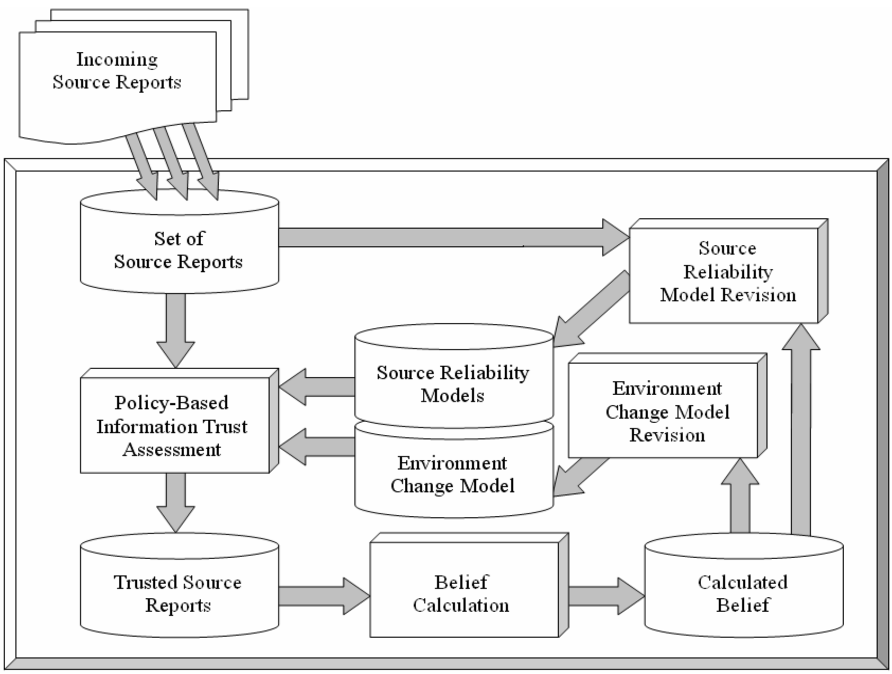

Fig. 2. Belief revision process. Information is valuated based on policies, then information most highly-valued is used to calculate the belief. Finally, models of source trustworthiness and truth change are updated

A high-level view of the belief revision process is illustrated in Figure 2. Each information source first sends an information report to the belief-maintaining agent. Next, the agent performs policy-based trust assessment of information reports to determine which reports should affect belief calculation. The belief distribution ϕ_B is then derived from these trusted reports; if the belief distribution is described by a normalized Gaussian, then the belief distribution mean, μ_B, and standard deviation, σ_B, are computed. The new belief is used to update models of source reliability and environment change patterns.

4.1 Information Trust Assessment Based on Information Recentness

In assessing the trustworthiness of information based on its recentness, we must first address the nature of information change over time. In a dynamic environment, once an agent has derived a belief about a belief subject v, the true value, x_T, of that belief subject can change over time. For some belief subjects (such as the location of an automobile), x_T might change very quickly, while for other belief subjects (such as the location of a building), x_T might change very slowly. In some domains, belief subjects might change so slowly that they are modeled as static.

If we assume that the agent's models of how x_T changes are non-Markovian (the agent does not maintain probability models to predict how v changes over time), then the agent can only assume that its information about v becomes less accurate over time. The rate at which this information loses accuracy is directly correlated to the rate at which x_T changes. For example, information about the location of an automobile might become obsolete very quickly, while information about the location of a building might remain relevant for much longer, since automobiles generally move more frequently than buildings do. [3] discuss independence and Markovian assumptions for fluents, which are "propositions whose truth value evolves over time."

The first step in developing an algorithm to assess the trustworthiness of information based on timeliness requires the agent to model the rate at which the belief subject, v, truly changes. Let δ_v represent a distribution describing the agent's approximation of the rate at which the true value, x_T, of v changes. Since the agent never knows the true value x_T of the belief subject, the agent must approximate δ_v based on its best approximations of x_T as contained in its past beliefs about v. More specifically, the agent must construct δ_v as a per-time-unit temporal error probability distribution based on the point values for the beliefs at times when belief intervals are instantiated. For this research, the agent can use a simple approach described here, assuming that the change in v is truly a consistent random-walk type of change: the agent calculates the change in the mean μ_B of its belief probability distribution ϕ_B at the beginning of each belief interval (when the belief is assumed most accurate), and divides that change by the (time) length of the belief interval. This unit of belief change for belief interval I_j, calculated as

$$\frac{\mu_{B,j+1} - \mu_{B,j}}{t_{j+1} - t_j}, \tag{2}$$

is included in the distribution δ_v, built up as numerous belief intervals are instantiated. We maintain a description of δ_v as a normal distribution with mean and standard deviation ($\mu_{\delta_v}, \sigma_{\delta_v}$).

Figure 3 demonstrates examples of the δ_v distribution for three cases: a) random movement with small change magnitude, b) random movement with large change magnitude, and c) movement with consistent direction and magnitude. δ_v provides a measure of the volatility of the belief's value over time. Since the trustworthiness of source information is determined based on the age of the information with respect to δ_v, then source information about highly volatile beliefs (δ_v is widely distributed or centered far from zero), become irrelevant very quickly, while values reported about more static beliefs (δ_v is narrowly distributed and centered at zero), remain relevant longer.

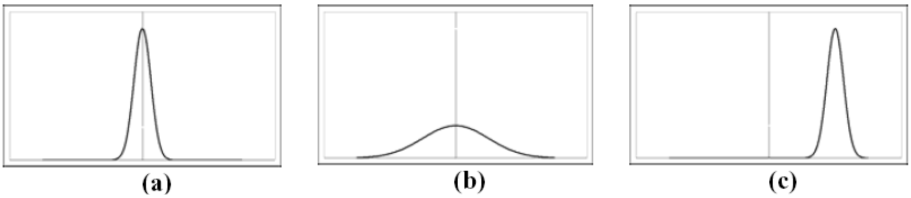

Fig. 3. Examples of δ_v change rate probability distributions, (a) δ_v describing a one-dimensional random movement with small change magnitude, (b) δ_v describing a one-dimensional random movement with large change magnitude, (c) δ_v describing a one-dimensional movement of consistent direction and magnitude

Since the only data used to estimate the change in the belief subject are beliefs located at discrete time points at which belief intervals are instantiated, it is possible that highly volatile (high frequency) cyclical change patterns will be incorrectly modeled as less volatile (lower frequency) cyclical change patterns, as explained by Nyquist's Theorem.[1] As experiments will show, even this simple estimator is a valuable indicator of the volatility of the belief value, expressing how the agent's valuation of the information should decrease as the data ages. In the future, calculations of δ_v might be based on more sophisticated function estimation and a combination of long-term and short-term trend analysis.

If the temporal error due to the passing of a single time unit can be modeled, then a three-dimensional temporal probability distribution can be constructed which extrapolates increasing information error due to belief subject change over time. When m is the number of timesteps describing the age of the information and μ_{δ_v} and σ_{δ_v} describe the per-timestep error distribution model of the information, the information error due to m passing timesteps can be described by a normal probability distribution with mean and standard deviation calculated as:

$$\mu_{\delta_v}(m) = m\mu_{\delta_v} \tag{3}$$

and

$$\sigma_{\delta_v}(m) = \sigma_{\delta_v}\sqrt{m} \tag{4}$$

.

[1] Nyquist's theorem states that a waveform may be reconstructed from samples taken at equal time intervals when the sampling rate is equal to, or greater than, twice the waveform frequency. Less frequent sampling rates result in a misconstructed waveform with a lower frequency.

Using these equations, the mean and variance of the error distribution are repeatedly summed for each passing timestep. In this paper we limit our scope to dealing only with random walk patterns for which we can assume μ_{δ_v} equals zero (as in Figures 3a and 3b). Incorporating μ_{δ_v} to describe how a belief changes over a belief interval results in an attempt at trend prediction, raising additional questions about belief uncertainty that are out of scope here. Therefore, for the cases we examine, σ_{δ_v} alone is a sufficient measure of temporal error.

4.2 Information Trust Assessment Based on Source Reliability

As described in previous research [1], an agent's model of information source reliability can be expressed as a belief, which allows use of existing representations. Additionally, the probability distribution within the belief representation allows multi-dimensional information about source reliability behavior to be captured; not only can the agent estimate the source's average reliability, but the source's consistency, or precision, as well. A simple method for constructing information source trust models is summarized below; for more detail, see [5].

Each source's reliability can be modeled as a distribution of errors between the source information distribution mean (μ_{s_i}) and the belief distribution mean (μ_B). Since the agent does not know the true value, x_T, of the belief subject, it can only evaluate the reliability of sources based on its belief, which is its best estimate of truth. To build this source reliability model distribution ρ_{s_i} reflecting the accuracy of the source s_i's distribution mean, after each timestep t in which a new report from the source is received, the error, α, between the source distribution mean and belief distribution mean, is calculated as:

$$\alpha(t) = \mu_{s_i}(t) - \mu_B(t) . \quad (5)$$

Over time, numerous $\alpha(t)$ point values build up the distribution ρ_{s_i}, described as a normal distribution by its mean, μ_{ρ,s_i}, and standard deviation, σ_{ρ,s_i}. For this research, we assume each source maintains static reliability characteristics; therefore, $\alpha(t)$ point values can be continually included to improve the accuracy of the trust model distribution. Future work will address the more common situation in which source reliability may change over time.

4.3 Instantiating a Belief Interval from Trusted Information

Once each type of possible error has been quantified, source information must be valuated according to the five information trust assessment policies, and a belief must be calculated. First, an error estimation probability distribution—encompassing error due to information age, source uncertainty, and source unreliability—can be calculated. The standard deviation $\sigma^*_{s_i}$ of the revised source distribution is found by summing the standard deviations for each type of error distribution:

$$\sigma_{s_i}^* = \sqrt{m\left[\sigma_{\delta_v}\right]^2 + \left[\sigma_{s_i}\right]^2 + \left[\sigma_{\rho,s_i}\right]^2} \tag{6}$$

The first term in the error summation represents error due to information age, and is intuitively based on both the age of the information (m timesteps) and the estimated belief subject rate of change (σ_{δ_v}). The second term represents the uncertainty communicated by the information source (σ_{s_i}), and the third represents the unreliability of the source itself (σ_{ρ,s_i}).

When the certainty of each source's information has been revised to account for these errors, contributing sources can be selected, based on the five information valuation policies, using a belief certainty maximization technique summarized here [5]. If the assumption holds that calculated belief standard deviation σ_B is an accurate measure of the closeness of the belief distribution mean μ_B to the true value x_T, then minimizing σ_B (thus maximizing certainty around μ_B) should yield a distribution most closely aggregated around x_T. Following from [5], the belief distribution standard deviation can be calculated from n sources as:

$$\sigma_B = \frac{\sqrt{\sum_{i=1}^{n}\left(\left(\sigma_{s_i}^*\right)^2 + \left(\mu_{s_i} - \mu_B\right)^2\right)}}{n}. \tag{7}$$

Minimizing σ_B addresses all five policies. First, source reliability, source certainty, and information age are accounted for in each source's revised report standard deviation ($\sigma_{s_i}^*$); smaller errors due to reliability, certainty, and age help to minimize σ_B. Second, source agreement is valued, since close distributions help to minimize σ_B. Finally, maximum information is encouraged, since larger number of n sources also helps to minimize σ_B.

The notion of source information trustworthiness, or relevance, can be explained in terms of the revised standard deviation of the source distribution. The information's relevance is directly proportional to the certainty of the information, implicitly described by the revised source distribution standard deviation $\sigma_{s_i}^*$. Using this concept, two source reports can be compared in terms of their relevance; one report from an unreliable source may be very recent, while another from a reliable source may be older. Assessing the relevance tradeoff between the two reports might be difficult, especially, since the relevance due to the data age is influenced by the change rate δ_v. In situations in which the belief value is relatively static, older data can more easily outweigh newer, less reliable data; in situations in which the belief value changes widely and often, all but the most recent data might be considered irrelevant. However, the method used to revise source distribution standard deviations quantifies these two relevance contributors (source reliability and information age) so they can be compared. Nonetheless, according to the belief certainty maximization method to be described below, though one source report might be deemed more relevant than another, both might be used to calculate the belief in order to satisfy the maximum information policy.

Relevant, contributing information is the subset of source reports which minimizes σ_B as calculated above (see [5] for a detailed discussion of optimizing the subset identification process). As seen from Equation 7, policy tradeoffs are assessed: the policy of utilizing maximum information is weighed against trusting only corroborated information or information deemed highly certain (where certainty is determined by source reliability, age of information, and certainty conveyed by the source). The subset of *n'* source reports resulting in the smallest calculated value for σ_B denotes the source information that should influence belief calculation. The belief distribution mean and standard deviation can be calculated from this subset using the following equations:

$$\mu_B = \frac{\sum_{i=1}^{n'} \mu_{s_i}}{n'} \tag{8}$$

$$\sigma_B = \frac{\sqrt{\sum_{i=1}^{n'} \left(\left(\sigma_{s_i}^* \right)^2 + \left(\mu_{s_i} - \mu_B \right)^2 \right)}}{n'} \tag{9}$$

Upon belief calculation, a new belief interval is instantiated with the calculated belief as the initial belief. The agent can then utilize the theory described in Section 4.1 to assess its certainty in its own belief over time. The probability distribution Φ_a^v can be calculated for continuous timesteps *m* after the instantiation of the new belief interval. The standard deviation of Φ_a^v is a function of time:

$$\sigma_B(m) = \sqrt{\left[\sigma_B\right]^2 + m\left[\sigma_{\delta_v}\right]^2} \tag{10}$$

The resulting increase over time in the probability distribution standard deviation can be termed *certainty depreciation*, since the agent devalues its belief over time by showing a decrease in certainty. For each step in which a change can occur, the probability distribution describing the total estimated belief error widens, so the certainty of one particular outcome decreases as many outcomes become equally plausible. Figure 4 shows belief intervals in which belief certainty is depreciated over time (compare to the belief intervals previously discussed in Figure 1, in which no certainty depreciation occurs). As shown from Equation 10, the rate of certainty depreciation is directly related to the agent's model (δ_v) of the rate at which *v* truly changes. For example, an agent might maintain a high certainty about the location of a building for longer that it might about the location of an automobile.

With the instantiation of a new belief interval, source reliability models and δ_v distributions can be recalculated. Revision of a source's reliability model only occurs at those belief intervals instantiated when information from that source is received. Otherwise, if the source's reliability is assessed based on out-of-date information, the source may be characterized as unreliable, when in reality, the information error is temporal, not due to source unreliability. Therefore, the information source's trust model is not penalized if the information was accurate when it was new, even if the information may currently be obsolete.

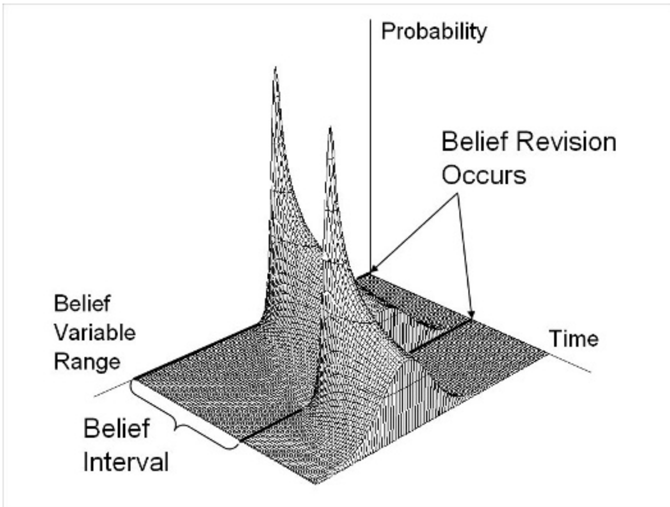

Fig. 4. Belief interval probability distributions demonstrating certainty depreciation

5 Experiments

Experiments are conducted in the domain of Unmanned Aerial Vehicle (UAV) target tracking to evaluate the effects of incorporating temporal error into the belief revision process. Therefore, the Temporal algorithm described previously in Section Four, which performs certainty depreciation and accounts for all policies, including the temporal policy, is compared against a Non-Temporal algorithm. The Non-Temporal algorithm does not perform certainty depreciation and does not account for temporal error (only maximum information, information corroboration, source certainty, and source reliability). The Non-Temporal algorithm is similar to the belief certainty maximization algorithm described in previous research [4]. The two experiment sets show the Temporal algorithm superior in the following aspects:

1. Incorporating a temporal policy in the belief revision process yields *more accurate beliefs* at the time of belief interval instantiation. By considering the age of information against other factors influencing information reliability, the beliefs resulting from this revision process will have more accurate means.
2. Practicing certainty depreciation across belief intervals (during periods between belief revision steps) yields *more accurate belief certainty assessments*. By decreasing belief certainty when no new information is available, the adjusted beliefs will have more accurate standard deviations.

In a single experiment run, a UAV agent is employed to monitor the location of a ground target. The target moves in a two-dimensional random walk, traveling a predetermined distance in each timestep. In a given timestep, the UAV agent receives reports from a subset of other UAVs, (here called source agents, for distinction), which simultaneously communicate estimates of the target's location to the UAV agent. Whether a source communicates a location estimate is dependent upon the reporting frequency specified for that source. The UAV agent uses source reports to calculate a belief about the true target location. Each source s_i reports estimated two-dimensional target location information as two (latitudinal and longitudinal) normal location probability distributions, $\phi_{s_i x}$ and $\phi_{s_i y}$, by specifying distribution means, $\mu_{s_i x}$ and $\mu_{s_i y}$, and standard deviations, $\sigma_{s_i x}$ and $\sigma_{s_i y}$.

In constructing experiments, each source is designed according to a ground truth reliability of either "reliable" or "unreliable." Reliability is defined by the source's accuracy in communicating target locations coordinates. Source s_i's location accuracy is described by a mean and standard deviation (μ_{rel,s_i}, σ_{rel,s_i}) describing a probability distribution of the source's error between its reported location distribution mean ($\mu_{s_i x}$ or $\mu_{s_i y}$) and the true target location. Given this definition, the mean of the error probability distribution corresponds to the accuracy of source location values to ground truth location values, and the standard deviation of the error probability distribution corresponds to the source's consistency, or precision of the source location values. Thus a source classified as "reliable" should have a mean error and/or a distribution standard deviation less than that of a source classified as "unreliable." To reflect the certainty in their information, all sources report a default, small standard deviation of 0.01.

For data collection purposes, only beliefs about one target's longitudinal ("x") coordinate are observed. Source reports are collected over experiments runs of 2000 timesteps each. The first experiment, assessing the importance of considering information age in forming accurate beliefs at revision time, examines only beliefs resulting from a belief revision (beliefs instantiating a new belief interval). In each experiment run, belief revision occurs approximately 200 to 400 times. During these timesteps, the Temporal algorithm evaluates source information based on several factors, including the age of the information, while the Non-Temporal algorithm does not consider information age. The second experiment considers only beliefs within a belief interval, excluding beliefs by which a belief interval is instantiated from new information. Each experiment run consists of approximately 1600 to 1800 timesteps in which belief revision is not performed. During these timesteps, the Temporal algorithm depreciates belief certainty, while the Non-Temporal algorithm leaves beliefs unchanged.

5.1 Experiment Parameters

The algorithms, Temporal and Non-Temporal, are assessed given the experiment parameters described in Table 1. Each experiment requires the specification of 1) number of information sources, n, 2) number of information sources, n_r, designated as "reliable", 3) number of sources, n_u, designated as "unreliable", 4) "reliable" source error specification, as given by a mean and standard deviation of the probability distribution of the source's error from ground truth (μ_r, σ_r), 5) "unreliable" source error specification, also given by mean and standard deviation of the source's error probability distribution (μ_u, σ_u), 6) "reliable" source reporting frequency, f_r, in terms of average number of timesteps between reports, and 7) "unreliable" source reporting frequency, f_u. The experiments vary the reporting frequency of reliable sources, creating a scenario in which accurate information from trustworthy sources is received less frequently than information from less reliable sources. Therefore, the receiving agent must evaluate the tradeoff between recent information from unreliable sources and older information from reliable sources. The number of reliable sources outweighs the number of unreliable sources, and unreliable sources report in every timestep.

Table 1. Experiment parameters. Both experiments vary reliable source reporting frequency, creating a scenario in which accurate information from trustworthy sources is received less frequently than information from less reliable sources

Number of Sources (n)	9
Number of Reliable Sources (n_r)	6
Number of Unreliable Sources (n_u)	3
Reliable Source Mean Error (μ_r)	0
Reliable Source Error Standard Deviation (σ_r)	0.5
Unreliable Source Mean Error (μ_u)	0
Unreliable Source Error Standard Deviation (σ_u)	2
Average Timesteps Between Reliable Source Reports (f_r)	{1,10,20,30,40,50,60,70,80}
Average Timesteps Between Unreliable Source Reports (f_u)	1

Two metrics are used to evaluate the algorithms [5]. The first metric evaluates belief accuracy by measuring mean error between the belief distribution mean, μ_B, and the ground truth value, x_T, averaged over m timesteps in which belief revision is performed:

$$\text{Error of Belief Distribution Mean} = \frac{\sum_{i=1}^{m} |\mu_B - x_{Ti}|}{m} . \tag{11}$$

The second metric assesses the accuracy of the belief's certainty assessment (described by the belief distribution standard deviation) by computing the fraction of the probability distribution ϕ_B assigned to an interval surrounding the ground truth value x_T. This metric assesses the probability assigned to a ground truth interval within the belief distribution ϕ_B:

$$\text{Truth interval probability} = \frac{\sum_{i=1}^{m} \left(\int_{x_T-\delta}^{x_T+\delta} \phi_B(x)dx \right)}{m} . \tag{12}$$

5.2 Experiment Results

Figures 5 and 6 show results for the first experiment, assessing the accuracy of beliefs *during* belief revision. As shown in Figure 5, the Temporal policy algorithm limits error of the belief distribution mean as the average number of timesteps between reliable source reports increases. However, the Non-Temporal algorithm maintains an increasing error of belief distribution mean. Therefore, the Temporal algorithm maintains significantly lower error than the Non-Temporal algorithm ($\alpha = 0.05$). Because the Temporal algorithm decreases trust in information as it ages, it is able to exclude information from reliable sources once that information becomes less accurate than recent information from unreliable sources. In contrast, the Non-Temporal algorithm consistently favors information from reliable sources, even if the information is inaccurate due to age. Figure 6 shows that the Temporal algorithm is able to maintain a slightly higher, but not significant (for $\alpha = 0.05$), truth interval probability than the Non-Temporal algorithm for most reliable reporting frequency variations. Thus the agent's ability to correctly assess its belief certainty does not change significantly sbetween algorithms.

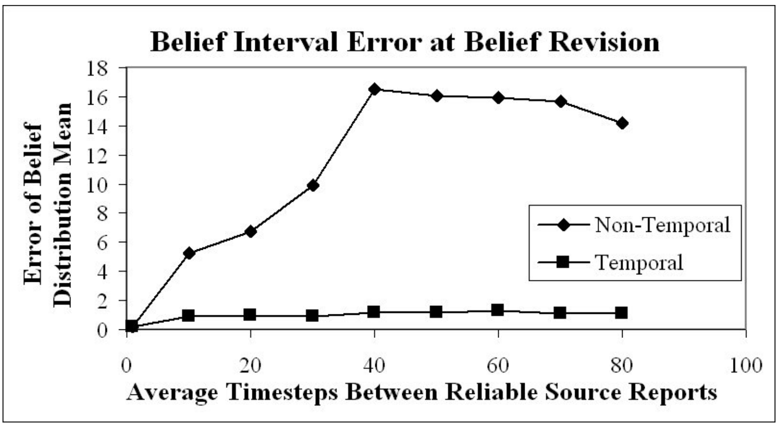

Fig. 5. Error of belief distribution mean vs. reliable source reporting frequency

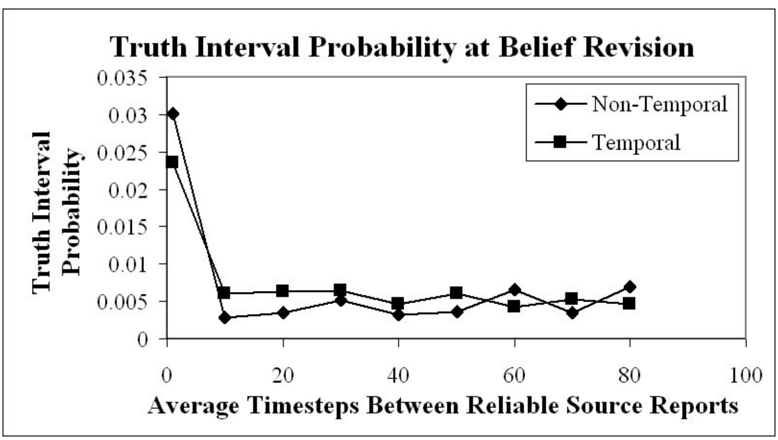

Fig. 6. Truth interval probability vs. reliable source reporting frequency

For the second experiment, assessing the accuracy of beliefs *between* belief revisions, results are shown in Figures 7 and 8. As shown in Figure 7, the error of belief means is only slightly lower for the Temporal algorithm. Beliefs derived using the Temporal algorithm benefit from the improved evaluation of source information at belief revision, as previously discussed in relation to Figure 5. However, belief means throughout the remainder of the belief interval are unchanged with both algorithms, yielding similar results for the Temporal and Non-Temporal algorithms. In addition, errors shown in Figure 7 are much higher than errors shown in Figure 5 because beliefs held several timesteps after belief revision are derived from increasingly old information. Importantly, Figure 8 demonstrates that the Temporal algorithm maintains significantly ($\alpha = 0.05$) higher truth interval probabilities that the

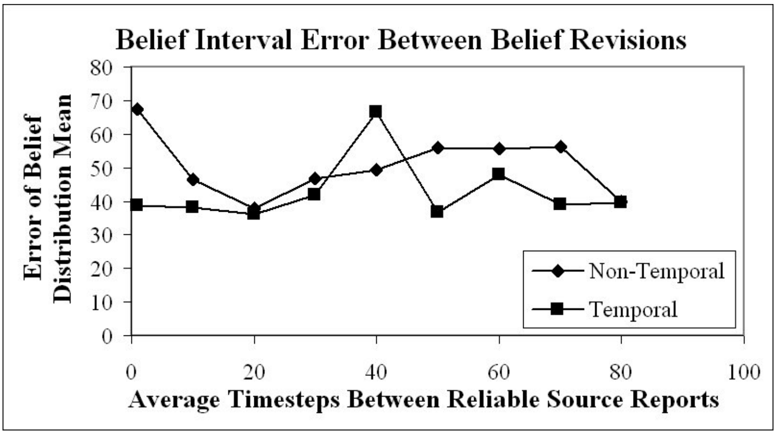

Fig. 7. Error of belief distribution mean vs. unreliable source error

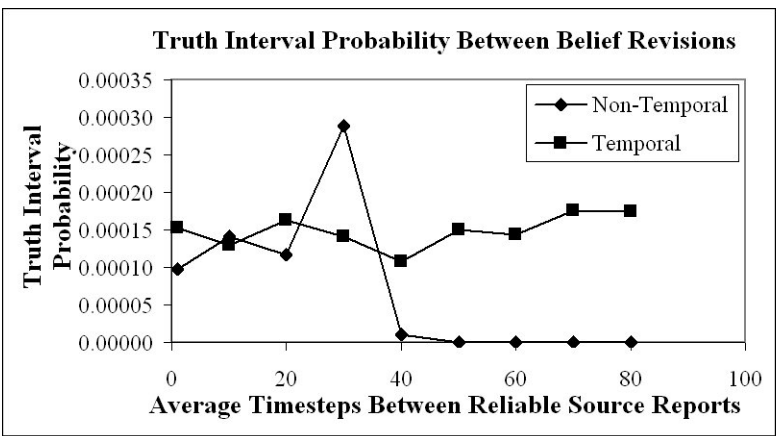

Fig. 8. Truth interval probability vs. unreliable source error

Non-Temporal algorithm for cases when reliable source report most infrequently. Belief certainty assessments for the Non-Temporal algorithm do not change over a belief interval and quickly become inaccurate. However, the Temporal algorithm is able to appropriately adjust belief certainty assessments as belief intervals age.

In summary, two algorithms are compared: a Temporal algorithm, which performs certainty depreciation and evaluates tradeoffs between all five policies described previously, and a Non-Temporal algorithm, which does not perform certainty depreciation and only addresses the remaining four policies. Results show that when a tradeoff between source reliability and information age must be evaluated, the Temporal algorithm is able to maintain significantly lower error of the belief distribution mean. In addition, the Temporal algorithm is able to maintain significantly higher truth interval probabilities, and thus more accurate certainty assessments, even as belief intervals age in the absence of new source information.

6 Conclusions

This research introduces a new policy for temporal information valuation, weighing the value of information, based on its recentness, against other evaluation criteria, such as the reliability of the source and the content of the information. The definition of a belief interval provides the agent with flexibility to acknowledge that a belief subject may be changing between belief revision instances. Since the belief interval framework describes the belief probability distribution over time, it allows for certainty depreciation in proportion to the belief subject rate of change.

In addition, this research presents a method for constructing the model of belief subject rate of change; this model is required if certainty depreciation is to represent increased error in a belief as time passes. Additionally, other causes of error, such as source unreliability and certainty conveyed by the source, are quantified against temporal error. Finally, the belief certainty maximization principle is used to select a subset of sources providing the highest quality of information, and beliefs are calculated from that contributing subset. Experimental results show the clear advantage of the Temporal algorithm, which performs certainty depreciation over belief intervals and evaluates source information based in information age. The Temporal algorithm derives more accurate beliefs at belief revision and maintains more accurate belief certainty assessments as belief intervals age than the Non-Temporal algorithm.

Assumptions made in this paper provide direction for future work. Belief subject rate of change is naively modeled as a normal distribution; in many cases, a Gaussian model permits distortion. This paper has limited the type of belief subject change examined; types of belief subject change other than random walks must be modeled, which may require predictive belief calculation. Other representations, and means of calculating belief subject rate of change, should be explored. As a further assumption, source reliability characteristics are modeled as static. In reality, information sources can change their behavior, perhaps gaining or losing competence or becoming malicious over time. The data used to build information source trust models must themselves be assessed for temporal relevance as sources change reliability characteristics.

Acknowledgment

This research is sponsored in part by the Defense Advanced Research Project Agency (DARPA) Taskable Agent Software Kit (TASK) program, F30602-00-2-0588. In addition, the research reported in this document/presentation was supported in part by the Countermeasures to Biological and Chemical Threats Program at the Institute for Advanced Technology of The University of Texas at Austin, under contract number DAAD17-01-D-0001 from the U.S. Army Research Laboratory. The views and conclusions contained in this document/presentation are those of the authors and should not be interpreted as presenting the official policies of the U.S. Government unless so designated by other authorized documents. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

References

1. Barber, K. S. and K. Fullam. (2003) "Applying Reputation Models to Continuous Belief Revision," In the Proceedings of *The Workshop on Deception, Fraud and Trust in Agent Societies at The Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2003),* Melbourne, Australia, July 14-18, pp. 6-15.
2. Barber, K. S. and J. Kim (2002). "Soft Security: Isolating Unreliable Agents," In the Proceedings of *The Workshop on Deception, Fraud, and Trust in Agent Societies at Autonomous Agents (AAMAS-2002)*, Bologna, Italy, July 15-19, pp. 8-17.
3. Dupin de Saint-Cyr, F. and J. Lang (1997). "Reasoning about Unpredicted Change and Explicit Time," In the Proceedings of *The First International Joint Conference on Qualitative and Quantitative Practical Reasoning*, Bad Honnef, Germany, June 9-12, pp. 223-236.
4. Fullam, K. and K. S. Barber (2004). "Using Policies for Information Valuation to Justify Beliefs," In the Proceedings of *The Third International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2004),* New York City, July 19-23.
5. Fullam, K. (2003). *An Expressive Belief Revision Framework Based on Information Valuation*, Master's Thesis, The University of Texas at Austin.
6. Jonker, C. M. and J. Treur (1999). "Formal Analysis of Models for the Dynamics of Trust Based on Experiences," In the Proceedings of *The 9th European Workshop on Modeling Autonomous Agents in a Multi-Agent World: Multi-Agent System Engineering (MAAMAW-99)*, Valencia, Spain, June 30-July 2, pp. 221-231.
7. Mui, L., A. Halberstadt, and M. Mohtashemi (2002). "Notions of Reputation in Multi-Agents Systems: A Review," In the Proceedings of *The First International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2002)*, Bologna, Italy, July 15-19, pp. 280-287.

8. Nodelman, U., C. Shelton, and D. Koller (2002). "Continuous Time Bayesian Networks," In the Proceedings of *The Eighteenth International Conference on Uncertainty in Artificial Intelligence,* Edmonton, Alberta, August 1-4, pp. 378-387.
9. Roorda, J., W. van der Hoek, and J. Meyer (2002). “Iterated Belief Change in Multi-Agent Systems,” In the Proceedings of *The First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2002),* Bologna, Italy, July 15-19, pp. 889-896.
10. Tawfik, A. and E. Neufeld (1994). “Temporal Bayesian Networks,” In the Proceedings of *The International Workshop on Temporal Reasoning*, Pensacola, Florida, May 4, pp. 85-92.
11. Yolum, P. and M. Singh (2003). "Achieving Trust via Service Graphs," In the Proceedings of *The Workshop on Deception, Fraud and Trust in Agent Societies at The Second International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS-2003),* Melbourne, Australia, July 14-18, pp. 76-83.