# Simulating Human Decision Making for Testing Soft and Hard/Soft Fusion Algorithms

Donald J Bucci, Sayandeep Acharya, and Moshe Kam
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, Pennsylvania, 19104-2816

***Abstract*—Current methods for evaluating the effects of human opinions in data fusion systems are often dependent on human testing (which is logistically hard and difficult to arrange for repeated tests of the same population). The alternative is to use hypothetical examples, which tend to be simplistic. To facilitate studies of data fusion architectures which integrate "soft" human-generated decisions, we have used a simulator of subjective beliefs. The simulator is based on the two-stage dynamic signal detection model of Pleskac and Busemeyer (2010). We use this scheme to simulate human opinions and combine them using belief fusion methods, including Bayes' Rule; Dempster's Rule of Combination (DRC); Yager's rule; the Proportional Conflict Redistribution Rule #5 (PCR5) from Dezert-Smarandache theory; and the consensus operator from subjective logic. In our simulations, the DRC and Bayes rule exhibited performance that was on par with, and in some cases better than PCR5 and the consensus operator (when used in conjunction with a measure of source reliability). In all simulated cases, Yager's rule exhibited inferior performance.**

## I. Introduction

The integration of subjective data sources in a data fusion system (also known as "soft fusion") has been studied intensely over the past few years [1]. In addition to the "hard fusion" of sensory information from devices that use electronic, optical, and acoustic modalities, there is a growing interest in augmenting decision making with available "soft" sensors (i.e., human opinions and assessments [1]). The incorporation of human opinions into data fusion systems could potentially improve accuracy and reliability. However, human opinions are difficult to model as they do not exhibit fixed error probabilities, and are often not easily characterized by probability distributions [1]. Furthermore, the employment of large numbers of humans for testing can be logistically challenging and expensive, and opportunities to re-test the same humans on modified data presentations and exposition schemes is often difficult. At least in the early stages of testing and tuning of data fusion algorithms, it may be desirable to use models of human decision making rather than using actual human-generated data.

Much work has been devoted over the past fifty years to developing new ways of representing and combining subjective and imprecise beliefs in areas such as pattern recognition, biometrics, medical diagnostics, and autonomous navigation. However, the majority of studies which include elements of human decision making have resorted to hypothetical examples, observing how fusion methodologies perform with respect to how a "logical and coherent human" would reason [2]–[4]. Other studies have adopted simple stochastic models, observing the performance of combined decisions through Monte Carlo simulations (e.g., confusion matrices [5]). There have also been several attempts to amass groups of humans for direct testing and analysis [6]. This last approach is in many ways preferred to the models it has replaced, but is logistically cumbersome and somewhat inflexible, especially in assessing systems and algorithms that require the tuning of a large number of parameters.

Models of human decision making from the social sciences and cognitive psychology have not been applied extensively to soft and hard/soft fusion systems. The few studies that have used such models tend to look at how task reward structures influence human decision-making strategies in situations where the human element acts in a supervisory role [7]. The present study seeks to analyze soft fusion systems where subjective data sources provide confidence assessments of the decisions they make. In Section II, we overview the model for human simulation used here, known as *two-stage dynamic signal detection* (2DSD). In 2DSD, humans are modeled via a tuple of parameters which direct a stochastic process that represents an internal evidence accumulation between two outcomes. The human tuples used here are taken from [8], and are the result of modeling human responses in a *line length discrimination task*, where human subjects are shown different pairs of lines and asked to identify (1) the longer of the two lines and (2) rate their confidence in their response on a subjective probability scale. In Section III, we review a few methods for representing and combining subjective beliefs. In Section IV, we describe ways of mathematically formulating human opinions for use with the belief combination methods of Section III. Finally, in Section V we use a subset of the data from [8] to simulate human opinions for the line length discrimination task and fuse them using the methods described in Sections III and IV.

## II. Human Simulation Methodology

### A. Two-Alternative Forced Choice Tasks

Human decision-making models have been a topic of interest for psychologists since the early 1960s [9]. The majority of work has been addressing decision making in *two-alternative forced choice (TAFC)* tasks, in which a subject is presented with a scenario and is forced to choose between two alternatives [9]. Models of decision making based on TAFC tasks

978-1-4673-5239-0/13/$31.00 ©2013 IEEE

assume that (1) internal evidence favoring each alternative is accumulated over time; (2) the process of internal evidence accumulation is subject to random fluctuations; and (3) a decision is made when a sufficient amount of internal evidence has been accumulated for one of the two alternatives [9]. As described in [8], many models of human decision making have been based on *Dynamic Signal Detection (DSD)*.

**Definition 1:** ***Dynamic Signal Detection (DSD) [8].*** *Consider a TAFC task given by the choice $A$ and its negation $\overline{A}$. Let the real-valued function $L(t)$ represent the accumulated evidence in favor of $A$*[1] *up until time instant $t$. A Dynamic Signal Detection (DSD) model is described by the stochastic linear difference equation*

$$\Delta L(t) = \delta \Delta t + \sqrt{\Delta t}\, \epsilon(t + \Delta t), \quad L(0) = L_0 \qquad \text{(II.1)}$$

*where $\delta \in \Re$ is known as the drift rate and $\epsilon(t)$ is a white noise process with zero mean and variance $\sigma^2$. The value $\sigma$ is known as the drift coefficient. The drift rate $\delta$ is either positive or negative, depending on whether $A$ or $\overline{A}$ is true. To make a choice, the process occurs until a threshold, either $\theta_A, \theta_{\overline{A}} \in \Re$ is crossed (where $-\theta_{\overline{A}} < L_0 < \theta_A$). The decision rule $\Lambda[L(t)]$ is given as*

$$\Lambda[L(t)] = \begin{cases} A & L(t) > \theta_A, \\ \overline{A} & L(t) < -\theta_{\overline{A}}, \\ \text{wait} & \text{otherwise.} \end{cases} \qquad \text{(II.2)}$$

By itself, DSD does not present a method for modeling subjective confidences. In *two-stage dynamic signal detection* (2DSD) [8], a DSD process is simulated until a decision is reached after time $t_d$. The evidence accumulation process of Equation II.1 is then continued for an additional $\tau$ seconds, after which the value of $L(t_d + \tau)$ is binned into a set of confidence intervals depending on what state was chosen at time $t_d$.

**Definition 2:** ***Two-Stage Dynamic Signal Detection (2DSD) [8].*** *Consider a DSD model describing an internal evidence accumulation $L(t)$ as defined for a TAFC as in Definition 1. Suppose a decision of $\omega \in \{A, \overline{A}\}$ is made at time $t_d$. Let $\mathbf{P}^{(\omega)} = [p_1^{(\omega)} \cdots p_{N_\omega}^{(\omega)}]$ denote the $N_\omega$ possible confidence values associated with choosing $\omega$. The assigned confidence level $p \in \mathbf{P}^{(\omega)}$ associated with deciding $\omega$ after waiting $t_c = t_d + \tau$ is given as*

$$p = p_i^{(\omega)} \quad \textit{when} \quad L(t_c) \in [c_{i-1}^{(\omega)}, c_i^{(\omega)}] \qquad \text{(II.3)}$$

*where $c_0^{(\omega)} = -\infty$ and $c_{N_\omega}^{(\omega)} = \infty$ for each $\omega \in \{A, \overline{A}\}$. The remaining confidence bin parameters $\mathbf{C}^{(\omega)} = [c_1^{(\omega)} \cdots c_{N_\omega - 1}^{(\omega)}]$ are chosen such that $c_{i-1} < c_i$ for each $i \in \{1, \ldots, N_\omega - 1\}$.*

Using Definitions 1 and 2, a human decision maker is represented via the $2(N_A + N_{\overline{A}}) + 4$ parameters as summarized in the tuple $\mathcal{S}$,

$$\mathcal{S} = \{\delta, \sigma, L_0, \theta_A, \theta_{\overline{A}}, \tau, \mathbf{C}^{(A)}, \mathbf{C}^{(\overline{A})}, \mathbf{P}^{(A)}, \mathbf{P}^{\overline{A}}\}. \qquad \text{(II.4)}$$

For a full description of these parameters, see [8, Table 2].

[1] $L(t) < 0$ represents evidence in favor of $\overline{A}$.

## B. 2DSD Simulator Implementation

The present study used data from [8] to simulate the fusion of opinions from six (6) subjects under a *line length discrimination task*. In the line length discrimination task described in [8], subjects were shown one of six possible pairs of horizontal lines. For any given pair, the subjects were asked to identify (1) which of the two lines was longer; and (2) rate their confidence in their response on the subjective probability scale $\{50\%, 60\%, \ldots, 100\%\}$. Six different line length pair types were presented, representing six different levels of difficulty. The work in [8] fit the subjects of the line length discrimination task using the 2DSD model parameters of Equation II.4 with the following additional restrictions: (1) the decision thresholds for both choices were set to be equal (i.e., $\theta_{\overline{A}} = \theta_A = \theta$); (2) the drift rate $\delta$ was chosen on a per trial (or simulation) basis from $N(\nu, \eta^2)$; (3) the initial condition $L_0$ was chosen on a per trial (or simulation) basis uniformly in the interval $[-0.5s_z, 0.5s_z]$, where $s_z \in \Re^+$ was the size of the interval; (4) the confidence bins for both choices were set equal (i.e., $\mathbf{C}^{(A)} = \mathbf{C}^{(\overline{A})} = \mathbf{C}$); (5) the confidence values were fixed for each subject as $\mathbf{P}^{(A)} = \mathbf{P}^{(\overline{A})} = [0.50, 0.60, \cdots, 1.00]$; and (6) the drift coefficient was fixed for each subject as $\sigma = 0.1$. Expression II.4 was thus reduced to the 10-tuple,

$$\mathcal{S} = \{\nu, \eta, s_z, \theta, \tau, c_1, c_2, c_3, c_4, c_5\}. \qquad \text{(II.5)}$$

The time step of the simulator was fixed at $\Delta t = 0.001$ for each subject. The parameter $\nu$ was allowed to vary with task difficulty. Specifically, the values $\nu_1$ and $\nu_6$ of [8, Table 6] represent the mean drift rates for a subject under the hardest and easiest task difficulties, respectively. The human simulator algorithm for a given $\mathcal{S}$ is shown in Figure 1. The parameter values used to simulate each subject can be found in [8, Tables 3 and 6]. Also, in [8, Table 6] separate decision thresholds $\theta$ were determined for two cases of the line length discrimination task: (1) when subjects were asked to focus on fast responses; and (2) when subjects were asked to focus on accurate responses. In the present study, the values of $\theta$ which represent the subjects focusing on accurate responses were used. For brevity's sake, the forthcoming simulation example (Section V) simulates opinions for each subject using the hardest task difficulty only ($\nu_1$).

# III. Belief Fusion Methodologies

## A. Bayesian Epistemology

Bayesian epistemology is the process of using the Kolmogorov axioms to represent logical constraints on degrees of belief, and probabilistic conditioning as the primary means of inferential reasoning [10], [11]. Let $\Omega$ represent a set of logical sentences which fully characterizes some phenomenon and $\mathcal{F}_\Omega$ represent a set algebra defined over $\Omega$. The probability function $P(\omega)$ represents the degree of belief a source holds towards the logical sentence $\omega \in \Omega$. For inferential reasoning between multiple belief assessments, a popular fusion method is *Bayes' rule*.

**Definition 3: Bayes' Rule.** *Suppose we observe a possible piece of evidence $x \in \mathcal{X}$ related to the logical sentences*

**Require:** $\mathcal{S}$ is a human source 10-tuple (expression II.5)
**function** 2DSDHUMANSIM($\mathcal{S}$)
    Choose $\boldsymbol{\delta}$ from $N(\nu, \eta^2)$
    Choose $\boldsymbol{L_0}$ from $U(-0.5s_z, 0.5s_z)$
    $\Delta t \leftarrow 0.001$,
    $\sigma \leftarrow 0.1$
    $t \leftarrow 0$
    $L \leftarrow \boldsymbol{L_0}$
    **while** $-\theta \leq L \leq \theta$ **do** ▷ Simulate a DSD process
        Choose $\boldsymbol{\epsilon}$ from $N(0, \sigma^2)$
        $\Delta L \leftarrow \boldsymbol{\delta}\Delta t + \boldsymbol{\epsilon}\sqrt{\Delta t}$
        $L \leftarrow L + \Delta L$
        $t \leftarrow t + \Delta t$
    **end while**
    **if** $L > \theta$ **then**
        $\omega \leftarrow A$ ▷ $A$ is the decided state
    **else if** $L < -\theta$ **then**
        $\omega \leftarrow \overline{A}$ ▷ $\overline{A}$ is the decided state
    **end if**
    $t_d \leftarrow t$
    $t_c \leftarrow t_d + \tau$
    **while** $t < t_c$ **do** ▷ Simulate interjudgment time
        Choose $\boldsymbol{\epsilon}$ from $N(0, \sigma^2)$
        $\Delta L \leftarrow \boldsymbol{\delta}\Delta t + \boldsymbol{\epsilon}\sqrt{\Delta t}$
        $L \leftarrow L + \Delta L$
        $t \leftarrow t + \Delta t$
    **end while**
    $c_0 \leftarrow -\infty$
    $c_6 \leftarrow \infty$
    $\mathbf{P} = [0.50, 0.60, 0.70, 0.80, 0.90, 1.00]$
    **for** $i = 1 \rightarrow 6$ **do** ▷ Confidence value binning
        **if** $c_{i-1} < L < c_i$ **then**
            $p \leftarrow p_i$ ▷ Extract the ith element from $\mathbf{P}$
        **end if**
    **end for**
    **return** $(\omega, p)$
**end function**

Fig. 1. Algorithm for implementing the 2DSD human simulator of [8] from the parameters given in Equation II.5.

*described by all $\omega \in \Omega$. Our original (a priori) beliefs, $P_\Omega(\omega)$, can be updated via Bayes' Rule:*

$$P_\Omega(\omega|x) = \frac{P_\mathcal{X}(x|\omega)P_\Omega(\omega)}{\sum_{\hat{\omega}\in\Omega} P_\mathcal{X}(x|\hat{\omega})P_\Omega(\hat{\omega})} \tag{III.1}$$

*where $P_\mathcal{X}(x|\omega)$ is known as the likelihood that the sentence $\omega \in \Omega$ would beget the evidence $x \in \mathcal{X}$. The quantity $P_\Omega(\omega|x)$ is known as our a posteriori belief in the logical sentence $\omega$, given our observation of the evidence $x$.*

## *B. Dempster-Shafer Theory*

The limitations of Bayesian epistemology are summarized in [10] and [11, Chapter 1]. One attempt to develop an alternative is given in *Dempster-Shafer (DS) Theory* [11]. Consider a set of logical sentences $\Omega$ and a corresponding set algebra $\mathcal{F}_\Omega$. For simplicity, we assume that $\Omega$ is finite and consists of disjoint elements. Hence, the algebra on $\Omega$ becomes the powerset $\mathcal{F}_\Omega = 2^\Omega$. In DS Theory, one is able to define one's beliefs across the powerset of $\Omega$ through a *belief mass assignment*.

**Definition 4: Belief Mass Assignment (BMA).** *A function $m : 2^\Omega \rightarrow [0, 1]$ is a belief mass assignment for some frame $\Omega$ if and only if $\sum_{X\subseteq\Omega} m(X) = 1$.*

As in [11], we assume that $m(\emptyset) = 0$. A BMA can be thought of as a normalized measure of the *explicit* evidence a source holds towards each subset of $\Omega$. As opposed to probability functions, BMAs can be used to specify evidence on a set of sentences $A \subseteq \Omega$ without imposing restrictions[2] on the individual sentences $\omega \in A$.

For a given set of sentences $A \subseteq \Omega$, it makes sense to define the *total amount of evidence assigned to* $A$ and the *amount of evidence which does not contradict* $A$. These quantities are known as *belief* and *plausibility* respectively.

**Definition 5: Belief/Plausibility Functions.** *Consider the BMA $m$ on some frame $\Omega$. The functions* $\text{Bel} : 2^\Omega \rightarrow [0, 1]$ *and* $\text{Pl} : 2^\Omega \rightarrow [0, 1]$ *are known as belief and plausibility functions on $\Omega$, and are given for any $A \subseteq \Omega$ as*

$$\text{Bel}(A) = \sum_{X\subseteq A} m(X) \tag{III.2}$$

*and*

$$\text{Pl}(A) = 1 - \text{Bel}(\overline{A}) = \sum_{X\cap A\neq\emptyset} m(X). \tag{III.3}$$

The difference $\text{Pl}(A) - \text{Bel}(A)$ describes the amount of *uncertain evidence* (or imprecise evidence) held towards $A$.

**Definition 6: Uncertainty Function**. *Consider the BMA $m$ on some frame $\Omega$. The uncertainty function* $\text{Un} : 2^\Omega \rightarrow [0, 1]$ *is given as*

$$\text{Un}(A) = \text{Pl}(A) - \text{Bel}(A) = \sum_{\substack{X\cap A\neq\emptyset \\ X\not\subseteq A}} m(X). \tag{III.4}$$

A few special types of belief functions are relevant to the present study. A *vacuous belief function* is one which represents total ignorance towards $\Omega$ such that

$$m(A) = \begin{cases} 1 & A = \Omega, \\ 0 & A \neq \Omega. \end{cases} \tag{III.5}$$

A *simple support function* assigns a degree of belief $s \in [0, 1]$ to some $A^* \subset \Omega$ such that

$$m(A) = \begin{cases} s & A = A^*, \\ 1 - s & A = \Omega, \\ 0 & \text{otherwise.} \end{cases} \tag{III.6}$$

It has been shown that Bayesian probabilities form a subset of belief functions. More specifically, a belief function whose corresponding BMA is non-zero for only the singleton elements of $2^\Omega$ is a Bayesian probability function [11, Theorem 2.8]. It is also easy to see from the definitions of belief

[2]For examples of these restrictions, see [11, Chapters 1-2]

and plausibility that $\text{Bel}(A) \leq \text{Pl}(A)$ for any $A \subseteq \Omega$, with equality only when $\text{Un}(A) = 0$. This relationship leads to a set of intervals $[\text{Bel}(A), \text{Pl}(A)]$ for every $A \subseteq \Omega$. These intervals beget a class of probability assignments, which can be thought of as *possible subjective probabilities* which the evidence supports[3].

Shafer cites *Dempster's Rule of Combination (DRC)* in [11] as the primary method for belief aggregation and updating.

**Definition 7: Dempster's Rule of Combination (DRC) [11].** *Consider the opinions of two sources, represented by the BMAs $m_1$ and $m_2$ respectively. Dempster's Rule of Combination is defined as an orthogonal combination of these belief functions. The resulting BMA $m_{1,2}(A)$ for any $A \subseteq \Omega$ is given by*

$$m_{1,2}(A) = \left(\frac{1}{1-\mathcal{K}}\right) \sum_{\substack{A_1, A_2 \subseteq \Omega \\ A_1 \cap A_2 = A}} m_1(A_1) m_2(A_2) \qquad \text{(III.7)}$$

*where $\mathcal{K} < 1$ is the measure of conflict after combination, given by*

$$\mathcal{K} = \sum_{\substack{A_1, A_2 \subseteq \Omega \\ A_1 \cap A_2 = \emptyset}} m_1(A_1) m_2(A_2). \qquad \text{(III.8)}$$

Since the DRC is both commutative and associative, extending it to the combination of more than two BMAs is straightforward.

Up to this point, DS Theory has not addressed how to take into account both source confidence and source reliability at the same time. Shafer suggests in [11, Chapter 11] a method for altering a BMA based on a level of perceived reliability known as *source discounting*. This notion was later redefined in [13] in terms of a BMA as follows.

**Definition 8:** ***Discounting on BMAs [13].*** *Let $\alpha \in [0,1]$ represent the reliability of a source. Source discounting the BMA $m$ for every $A \subseteq \Omega$ is given by*

$$m(A;\alpha) = \begin{cases} \alpha m(A) & A \neq \Omega, \\ \alpha m(A) + (1-\alpha) & A = \Omega. \end{cases} \qquad \text{(III.9)}$$

Source discounting can also be performed for Bayesian probabilities as follows.

**Definition 9:** ***Discounting on Bayesian Probabilities.*** *Let $\alpha \in [0,1]$ represent the reliability of a source. Source discounting the Bayesian probabilities $P$ for every $\omega \in \Omega$ is given by*

$$P(\omega;\alpha) = \alpha P(\omega) + |\Omega|^{-1}(1-\alpha) \qquad \text{(III.10)}$$

*where $|\Omega|$ represents the cardinality of $\Omega$.*

Source discounting for Bayesian probabilities can be thought of as increasing the level of entropy in the source's belief proportional to $\alpha$. For a fully discounted Bayesian probability, all outcomes are made equiprobable (the analogue of the vacuous BMA in a DS theoretic construction).

[3]Probabilities defined in this way are referred to as "probabilities of provability" and in general do not carry a statistical interpretation [12]

### *C. Variations on Dempster-Shafer Theory*

If the combined sources present highly conflicting opinions and it is not feasible to perform source discounting, a DS theoretic approach may not provide an acceptable result [2], [14]. One variation, *Yager's Rule*, calls for interpreting the evidence in conflict as uncertain evidence.

**Definition 10:** ***Yager's Rule [15].*** *Consider two BMAs $m_1$ and $m_2$ over some set of logical sentences $\Omega$. The resulting BMA $m_{1,2}$ after combination via Yager's rule for any $A \subseteq \Omega$ is given as*

$$m_{1,2}(A) = \begin{cases} \sum_{\substack{A_1, A_2 \subseteq \Omega \\ A_1 \cap A_2 = A}} m_1(A_1) m_2(A_2) & A \neq \Omega \\ m_1(A) m_2(A) + \mathcal{K} & A = \Omega \end{cases} \qquad \text{(III.11)}$$

*where $\mathcal{K}$ is the degree of conflict between $m_1$ and $m_2$ as given in expression III.8.*

Yager's rule is commutative but in general not associative. It is also possible to subdivide evidence amongst conflicting propositions proportionate to the degrees of belief held by the sources before combination. This concept has lead to the family of *Proportional Conflict Redistribution (PCR)* rules by Dezert and Smarandache. In this paper we investigate *PCR5*, the PCR rule most widely used at present [16].

**Definition 11:** ***Proportional Conflict Redistribution Rule #5 (PCR5) [16].*** *Consider two BMAs $m_1$ and $m_2$ over some set of logical sentences $\Omega$. The resulting BMA $m_{1,2}$ after combination via PCR5 for any $A \subseteq \Omega$ is given as*

$$\begin{aligned} m_{1,2}(A) = & \sum_{\substack{A_1, A_2 \subseteq \Omega \\ A_1 \cap A_2 = A}} m_1(A_1) m_2(A_2) + \\ & \sum_{\substack{X \subseteq \Omega \\ X \cap A = \emptyset}} \left( \frac{m_1(X)^2 m_2(A)}{m_1(X) + m_2(A)} + \right. \\ & \left. \frac{m_2(X)^2 m_1(A)}{m_2(X) + m_1(A)} \right). \end{aligned} \qquad \text{(III.12)}$$

Similar to Yager's Rule, the PCR5 is commutative but in general not associative [16]. Finally, subjective logic [17] uses the general terminology of BMAs, belief functions, and uncertainty while providing a way of representing any $\Omega$ as a binary frame focused on some $A \subseteq \Omega$ and its negation. The result is a four-valued vector known as an *opinion tuple*, which can be rephrased in terms of BMAs for a binary $\Omega$ as follows.

**Definition 12:** ***Consensus Operator, Binary Frames [17].*** *Consider the binary frame $\Omega = \{A, \overline{A}\}$. The consensus operator $m_{1,2}$ can be represented completely in terms of the BMAs $m_1$ and $m_2$ as*

$$\begin{aligned} m_{1,2}(A) &= \frac{m_1(A) m_2(A \cup \overline{A}) + m_2(A) m_1(A \cup \overline{A})}{\mathcal{K}_C} \\ m_{1,2}(\overline{A}) &= \frac{m_1(\overline{A}) m_2(A \cup \overline{A}) + m_2(\overline{A}) m_1(A \cup \overline{A})}{\mathcal{K}_C} \\ m_{1,2}(A \cup \overline{A}) &= \frac{m_1(A \cup \overline{A}) m_2(A \cup \overline{A})}{\mathcal{K}_C} \end{aligned} \qquad \text{(III.13)}$$

*where* $\mathcal{K}_C = m_1(A\cup\overline{A})+m_2(A\cup\overline{A})-m_1(A\cup\overline{A})m_2(A\cup\overline{A})$. *If* $\mathcal{K}_C = 0$, *then the consensus operator becomes*

$$\begin{aligned} m_{1,2}(A) &= \frac{m_1(A)+m_2(A)}{2}, \\ m_{1,2}(\overline{A}) &= \frac{m_1(\overline{A})+m_2(\overline{A})}{2}, \\ m_{1,2}(A\cup\overline{A}) &= 0. \end{aligned} \quad \text{(III.14)}$$

The consensus operator is commutative and associative [17].

## IV. BMA Formulation Methods

According to the algorithm of Figure 1, a simulated human opinion produces a decision $\omega \in \{A, \overline{A}\}$ and a confidence value $p \in [0,1]$. To use the combination operators described in Section III, $\omega$ and $p$ must be translated into BMAs (or in the case of Bayes' Rule, probability functions). In the present study, $\omega$ and $p$ were used to formulate simple support functions of degree $s$ focused on $\omega$, as in expression III.6. For Bayes' rule, the likelihood values of Equation III.1 were formed such that $P_{\mathcal{X}}(x|\omega) = s$ and $P_{\mathcal{X}}(x|\overline{\omega}) = 1-s$. The degree of support $s$ was set in the following four ways

1) **Decisions Only:** Each subject's source strength $s$ was estimated beforehand by determining the subject's proportion of correct decisions over 2,000 simulations. The subject's simulated confidence value $p$ was unused.
2) **Confidences Only:** Each subject's source strength $s$ was taken as the simulated confidence value $p$ produced at each run of the human simulator.
3) **Decision Reliability Discounting:** Each subject's source strength $s$ was taken as the simulated confidence value $p$ produced at each run of the human simulator. The resulting simple support function was then discounted using expression III.9 (or in the Bayesian case, expression III.10). The discount rate $\alpha$ was taken as the subject's proportion of correct decisions as described in the "Decisions Only" BMA formulation method.
4) **Evidence Strength Discounting:** Each subject's source strength $s$ was taken as the simulated confidence value $p$ produced at each run of the human simulator. The resulting simple support function was then discounted using expression III.9 (or in the Bayesian case, expression III.10). The discount rate $\alpha$ was taken as the subject's average evidence strength, estimated beforehand over 2,000 simulations. The evidence strength $\mathcal{E}(\omega, p)$ of a subject's decision $\omega$ and confidence assessment $p$ was given as

$$\mathcal{E}(\omega, p) = 1 - BS(\omega, p|\omega^*), \quad \text{(IV.1)}$$

where $\omega^*$ is the true outcome and $BS(\omega, p|\omega^*)$ is the quadratic scoring rule known as the Brier score [8].

**Definition 13:** ***Brier Score [8].*** *Let* $p \in [0,1]$ *be a confidence assessment towards the outcome* $\omega \in \{A, \overline{A}\}$. *Let* $\omega^* \in \{A, \overline{A}\}$ *represent the true outcome. The Brier score is a measure of distance on the interval* $[0,1]$ *between the given confidence assessment and a confidence assessment which assigns the true outcome probability one, given by*

$$BS(\omega, p|\omega^*) = \begin{cases} (1-p)^2 & \omega = \omega^*, \\ p^2 & \omega \neq \omega^*. \end{cases} \quad \text{(IV.2)}$$

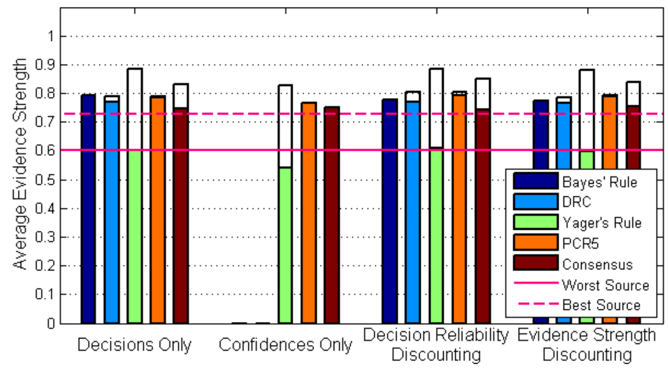
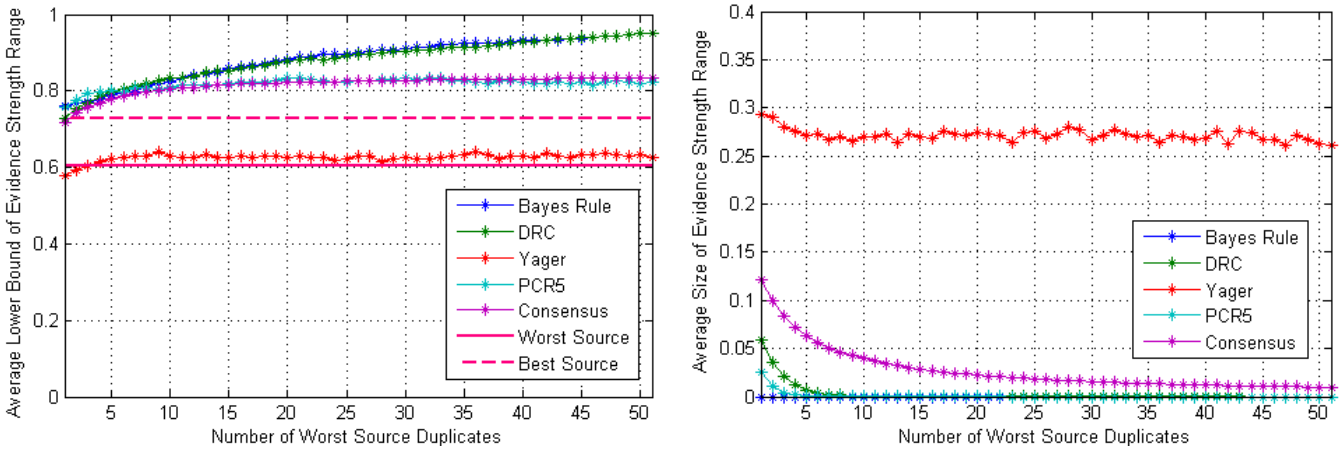



Fig. 2. Evidence strengths of fused subjective data on the line length task, hardest difficulty. Clear bars indicate the evidence strength ranges for each fusion method. Higher evidence strengths and smaller evidence strength ranges are better.

## V. Simulations

The BMA formulation methods given in Section IV and the combination operators described in Section III were used to perform fusion on the simulated subjects of [8]. All belief combination methods were initialized using vacuous BMAs as defined in expression III.5 (or in the case of Bayes' rule, equiprobable outcomes). All BMA formulation methods were simulated for each combination method over 2,000 runs. Performance for Bayesian case was taken as the average combination evidence strength $\mathcal{E}(\omega^*, p_{Bayes}|\omega^*)$, where $p_{Bayes}$ was the resulting *a posteriori* probability after combination via Bayes' rule. For the belief function combination methods, ranges on combination evidence strength were generated using belief and plausibility values as lower and upper bounds on possible subjective probabilities. The size of the evidence strength range represents the *precision* of the combination rule, and the lower bound of the evidence strength range represents the *accuracy* of the combination rule. An ideal combination rule should have high accuracy (i.e., evidence strength close to one) and high precision (i.e., a small evidence strength range).

Figure 2 shows the average evidence strength of each combination method (namely, Bayes' Rule, DRC, Yager's Rule, PCR5 and the Consensus operator), across the four BMA formulation methods discussed in Section IV (namely, Decisions Only, Confidences Only, Decision Reliability Discounting, and Evidence Strength Discounting). For comparison, the highest and lowest average evidence strengths among the six subjects are also shown. Subjects were combined in groups of two in the order $(((\mathcal{S}_1 \odot \mathcal{S}_2) \odot \mathcal{S}_3) \odot \cdots) \odot \mathcal{S}_6)$, where the symbol $\odot$ represents any of the combination rules from Section III, and the subscript of $\mathcal{S}$ represents the subject number as given in the columns of [8, Table 6]. Both Bayes' rule and the DRC were unable to combine subjective confidences consistently without performing source discounting. In all other cases, Bayes' rule and the DRC produced combination evidence strengths on par with the PCR5 and the consensus operator. Yager's rule was found to produce the lowest combination evidence strengths and the largest combination evidence strength ranges. It was also observed that the BMA formulation methods which used source discounting produced slightly higher combination




(a) (b)

Fig. 3. Evidence strengths of combined subjective confidences, discounted by evidence strength in the line length task. Results shown for the hardest difficulty. Duplicates of the subject which exhibited the lowest evidence strength included in the combination. (a) Average lower bound of the evidence strength range versus the number of worst source duplicates present in the combination. Higher evidence strengths are better. (b) Size of evidence strength range versus the number of worst source duplicates present in the combination. Smaller evidence strength range sizes are better.

evidence strengths than the ones which did not. Discounting by decision reliabilities and source evidence strengths produced similar combination evidence strengths.

Figure 3 shows the average strength of each combination method (namely, Bayes' Rule, DRC, Yager's Rule, PCR5 and the Consensus operator) when discounting the sources using their estimated evidence strengths, and while also including an increasingly large number of duplicates of the subject which exhibited the lowest evidence strength in the combination (i.e., the "worst" source). These subjects were combined in a similar manner as in Figure 2, except that the worst source and its duplicates were combined after the other five sources. When more than 10 duplicates of the worst source were included in the combination, Bayes' rule and the DRC produced higher combination evidence strengths than the other combination methods (Figure 3a). The size of the evidence strength range was found to decrease the fastest with PCR5 and the slowest with the consensus operator (Figure 3b). The size of the evidence strength range for Yager's rule was found to stay relatively constant.

## VI. Conclusions

We have demonstrated how it is possible to use a model of human decision making and confidence assessment to construct a simulator for evaluating the performance of data fusion systems using subjective opinions. In our simulations, Bayes' Rule, the DRC, the PCR5, and the consensus operator performed similarly when source discounting was performed. Both Bayes' rule and the DRC were unable to combine subjective confidences consistently when source discounting was not performed. Finally, when including more than 10 duplicates of an inferior-quality source in the combination, Bayes' rule and the DRC produced higher evidence strengths than the other combination methods. Future work includes extending the simulator for use on $M$-ary ($M > 2$) $\Omega$, and eventually the development of a test bed for hard and soft data fusion algorithms for multiple tasks and models of human decision makers.

## References


[1] D. L. Hall, M. McNeese, J. Llinas, and T. Mullen, "A framework for dynamic hard/soft fusion," in *Information Fusion*, pp. 1–8, 2008.

[2] L. A. Zadeh, "A simple view of the dempster-shafer theory of evidence and its implication for the rule of combination," *AI Magazine*, vol. 7, no. 2, pp. 85–90, 1986.

[3] K. Premaratne, M. Murthi, J. Zhang, M. Scheutz, and P. Bauer, "A dempster-shafer theoretic conditional approach to evidence updating for fusion of hard and soft data," in *Information Fusion, 2009. FUSION '09. 12th International Conference on*, pp. 2122–2129, July 2009.

[4] A. Josang, "A logic for uncertain probabilities," *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 9, pp. 279–311, 2001.

[5] D. Buede and P. Girardi, "A target identification comparison of bayesian and dempster-shafer multisensor fusion," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 27, pp. 569–577, September 1997.

[6] S. Acharya and M. Kam, "Evidence combination for hard and soft sensor data fusion," in *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pp. 1–8, July 2011.

[7] M. Cao, A. Stewart, and N. Leonard, "Convergence in human decision-making dynamics," *Systems & Control Letters*, vol. 59, no. 2, pp. 87–97, 2010.

[8] T. Pleskac and J. Busemeyer, "Two-stage dynamic signal detection: A theory of choice, decision time, and confidence," *Psychological review*, vol. 117, pp. 864–901, July 2010.

[9] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. Cohen, "The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks.," *Psychological review*, vol. 113, no. 4, pp. 700–765, 2006.

[10] W. Talbott, "Bayesian epistemology," in *The Stanford Encyclopedia of Philosophy* (E. N. Zalta, ed.), Stanford University, 2011.

[11] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press: Princeton, NJ, 1976.

[12] J. Pearl, "Reasoning with belief functions: An analysis of compatibility," *International Journal of Approximate Reasoning*, vol. 4, no. 5-6, pp. 363–389, 1990.

[13] P. Smets, "Decision making in a context where uncertainty is represented by belief functions," in *Belief Functions in Business Decisions* (R. P. Srivastava and T. J. Mock, eds.), pp. 17–61, Physica-Verlag, 2002.

[14] J. Dezert, P. Wang, A. Tchamova, *et al.*, "On the validity of dempster-shafer theory," in *Fusion 2012-15th International Conference on Information Fusion*, 2012.

[15] R. Yager, "On the dempster-shafer framework and new combination rules," *Information Sciences*, vol. 41, pp. 93–138, 1987.

[16] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion*, vol. III. American Research Press, 2009.

[17] A. Josang, "The consensus operator for combining beliefs," *Artificial Intelligence*, vol. 142, pp. 157–170, 2002.