UNCLASSIFIED

Australian Government

Department of Defence

Defence Science and Technology Group

Influx$_1$: A Tool and Framework for Reasoning under Uncertainty

Van Nguyen and Michael Docking

Cyber and Electronic Warfare Division
Defence Science and Technology Group

DST Group–TR–3142

ABSTRACT

Influx is a reasoning component of REALISE, a generic platform for intelligent systems currently under investigation as part of an ongoing research program conducted by the CAO Branch. Influx is envisioned to have two major uses: facilitating and mediating reasoning processes for intelligent missions, and serving as a tool and framework for reasoning under uncertainty. This document reports on some initial research and development efforts pertaining to the reasoning aspect of Influx in the latter scenario. Due in part to its generality, the Dempster-Shafer (D-S) theory is chosen as a theoretical basis for representing imperfect knowledge and for reasoning with such knowledge. Since Influx is intended to deal with real-time and real-life applications, it is of primary importance for the reasoning tool to practically achieve adequate performance, flexibility, scalability and system dynamics. To this end, Influx aspires to reach such objectives while attaining a plausible reasoning mechanism formulated from D-S methods and techniques. The initial version, Influx$_1$, is a simple, but highly efficient and flexible, nonmonotonic rule-based system enhanced with D-S based belief representation, fusion and inference. Influx$_1$ has been applied towards tasks that include situational awareness and network traffic analysis. This document provides a high-level description of Influx$_1$ from the reasoning perspective. Research and development pertaining to the implementation of the reasoning tool and specific applications are not included in this document.

RELEASE LIMITATION

Approved for Public Release

Published by

Cyber and Electronic Warfare Division
Defence Science and Technology Group
PO Box 1500
Edinburgh, South Australia 5111, Australia

Telephone: 1300 333 362
Facsimile: (08) 7389 6567

AR-016-372
September, 2015

APPROVED FOR PUBLIC RELEASE

Influx$_1$: A Tool and Framework for Reasoning under Uncertainty

Executive Summary

Influx is a reasoning component of a larger platform for intelligent systems currently under investigation as part of an ongoing research program conducted by the CAO Branch of CEWD, Defence Science and Technology Group. This document reports on the initial version Influx$_1$, a tool and framework currently used for developing and executing certain reasoning tasks under uncertainty. The exposition given throughout the document is purely focused on the reasoning perspective.

Since *uncertainty* is a major notion associated with knowledge representation and reasoning in Influx, the report begins with an overview of the major paradigms (namely, numerical and symbolic) for uncertainty representation, which is followed by an identification of reasons for adopting a numerical approach on the ground of efficiency and the accuracy of uncertainty measurement. To further clarify the types of uncertainty and the specific formalisms suited to Influx$_1$, the document gives a brief introduction and comparison of the three major theories (specifically, probability theory, possibility theory and the Dempster-Shafer or D-S theory) which are at the foundation of many such quantitative reasoning approaches. The comparison allows D-S theory to be chosen as a theoretical basis for the reasoning framework due to its generality.

In order to assist the reader in comprehending the various reasoning aspects of Influx$_1$ to be discussed, the report provides the necessary background for D-S theory, clarifying common practical concerns (such as computational complexity, evidence independence assumption, conflict handling and the reliability of sources), and discussing how these concerns have been addressed in the literature. This is followed by a high-level description of the reasoning framework in Influx$_1$ according to the following major perspectives: *concept representation*, *belief representation*, *belief combination* and *belief propagation*.

Regarding concept representation, the report gives an explanation of how knowledge can be represented in Influx$_1$ and how a complex piece of knowledge can be constructed based on simpler ones. As the reasoner aims to support real-world applications, a piece of knowledge in Influx$_1$ is often not known with certainty, but associated with a belief capturing the degree of uncertainty associated with the piece of knowledge. In this regard, the document discusses the multiple methods for belief representation in Influx$_1$, allowing for a rich and informative interpretation of knowledge. As it is possible that there is more than one belief associated with a piece of knowledge (which is usually the case when beliefs are provided by multiple sources or induced by a collection of evidence), there is a need to combine such beliefs. In this regard, the document presents a collection of methods for belief combination implemented in Influx$_1$. Since such belief combination methods possess

different characteristics which can prevail in certain scenarios, a broad guideline regarding how the methods can be applied in various practical contexts is also given. It is also often the case that only a small portion of the knowledge in the knowledge base has access, or is visible, to the external world, and thus can receive evidence from dedicated sources. The majority of the knowledge would receive no direct observations, and therefore assessment of their belief must be inferred (if possible) from other pieces of knowledge. To this end, the report explicates how the relationship between any two pieces of knowledge is formulated in the framework of D-S theory, and presents a collection of belief propagation functions that allow one to infer the degree of belief for knowledge which the collected evidence does not directly bear on.

Finally, a summary of the merits and limitations of the reasoning framework with respect to flexibility, efficiency, scalability and system dynamics is given.

Authors

Van Nguyen
Cyber and Electronic Warfare Division

Van Nguyen joined DST Group in 2010 as a research scientist. Her research focus is on artificial intelligence and its application to computer security problems. This has included research in the areas of reasoning under uncertainty, attribution of spear phishing attacks, and ontology and knowledge representation. Van obtained her PhD in 2010 with a thesis on the implementation of membrane computing models on reconfigurable hardware.

Michael Docking
Cyber and Electronic Warfare Division

Michael Docking joined DST Group in 2000 as a computer systems engineer, developing hardware, software and architectures for cyber security, intelligent systems and computing. Michael leads the Hinder project for autonomous cyber defence, incorporating Influx.

THIS PAGE IS INTENTIONALLY BLANK

Contents

1 Introduction

Primarily designed as a knowledge and reasoning backbone of a high-level and distributed reasoning platform, Influx is not intended to provide sophisticated reasoning. Instead, this reasoning tool aims to offer core reasoning capabilities with imperfect knowledge that can potentially satisfy many practical requirements, including reasoning in real-time, handling a potentially very large amount of data, flexibly supporting different reasoning modalities, and supporting dynamic changes and updates in the knowledge structure. Advanced and complex reasoning tasks, if desired, would be built on top of Influx, or delegated to dedicated tools integrated with the reasoning system.

To open the possibility for Influx to flexibly handle and reason with knowledge with different natures of 'imperfectness', Dempster-Shafer (D-S) theory is adopted as a theoretical basis for the design of Influx. This line of exploration and investigation has resulted in the development of two versions of Influx: Influx_1 and Influx_2.

Influx_1 is a simple, but highly flexible and efficient, tool and framework for reasoning under uncertainty. The high flexibility of Influx_1 is due to the utilisation of various methods and techniques pertaining to D-S belief representation, fusion and inference. The high efficiency of Influx_1 partly owes to the simplified reasoning mechanism, making it possible to devise highly-optimised techniques for data storage and algorithm execution at the implementation level. Influx_1 has been applied towards tasks that include situational awareness and network traffic analysis.

Influx_2 is motivated by the requirement to enhance the reasoning capability of Influx_1 while attempting to minimise as much as possible the amount of complexity added to the reasoning tool, and thus retain the practicability of Influx_1. To this end, Influx_2 capitalises on methods and techniques in the paradigm of evidential networks so as to facilitate deductive and abductive reasoning in a seamless manner and to rectify certain restrictions of Influx_1. Nevertheless, evidential networks and related models, while possessing a mathematically sound and powerful mechanism to represent and reason with imperfect knowledge, often bear high computational complexity and memory consumption. Many of them also require a static and controlled execution environment, rendering themselves impracticable in many real-life situations. This has led us to implement a simplified and customised form of an evidential network in Influx_2 which potentially offers a significant decrease in space and time consumption, and potentially promotes system dynamics. As a result, we have devised nonconventional and approximate representation and reasoning techniques where conventional and exact methods are no longer possible due to the various restrictions imposed on the network structure.

An exposition of Influx_2 will be given in another document. This document provides a high-level description of Influx_1 from the reasoning perspective. The organisation of the document is given below.

Section 2 presents a brief discussion pertaining to the formal representation of uncertainty, with a focus on numerical formalisms. The section also clarifies the types of uncertainty to be addressed in Influx, and presents the rationale for choosing D-S theory as the theoretical basis for the reasoning system.

Section 3 provides an overview of the applied aspects of D-S theory in the particular context of information fusion. In addition, the section discusses and addresses major concerns in the literature pertaining to the application of D-S theory in a practical context.

Section 4 presents a high-level description of Influx_1 from the reasoning perspective, detailing how D-S theory serves as a basis for uncertainty representation and reasoning in the system.

Section 5 concludes the document with a summary of the merits and limitations of Influx_1.

In order to make the material accessible to a wide range of audiences, the content of this report is presented in a conceptual and intuitive manner.

2 Uncertainty representation

The real world is characterised by uncertainty and constantly changing conditions. Any reasoning system that strives to support real-world applications should be able to appropriately handle uncertain information. In this regard, a consideration of the relevant types of uncertainty, and subsequently of the appropriate formalism to represent and reason with uncertain knowledge is a prerequisite to the design and development of any practical reasoning system.

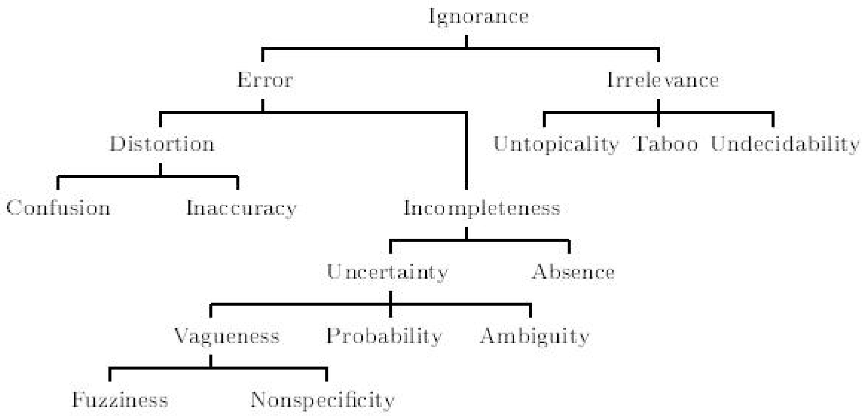


Figure 1: A taxonomy of ignorance, providing one interpretation of uncertainty [52] (cited in [38]). In this taxonomy, uncertainty is viewed as one type of incompleteness within the broad spectrum of 'ignorance'.

Formal representations of uncertainty[1] have drawn substantial research interest, giving rise to various interpretations and formulations. One such interpretation is illustrated in Figure 1, while others can be consulted at [56, 60, 63]. The different ways to describe uncertainty inevitably lead to efforts to devise corresponding modelling and computational techniques — the multitude of diverse formalisms that result can be categorised as *numerical* and *symbolic*.

In numerical approaches, certainty is quantitatively measured by a numerical value, whether it is a probability measure (as in probability theory), possibility and necessity measures (as in possibility theory), or certainty factors (as in rule-based systems). Conversely, in (purely) symbolic approaches, uncertainty is handled by non-numerical techniques. More specifically in the framework of modal logics, uncertainty is qualified by symbolic modalities (e.g., *possible* and *necessary*), and is expressed by means of *relative confidence relations between propositions*, as opposed to numerical evaluations [14]. In other non-monotonic formalisms (such as default logics [43] and McCarthy's circumscription [34]), uncertainty

[1] Please note that in the literature, the term *uncertainty* is both used to (i) broadly capture the whole spectrum of possible definitions and interpretations pertaining to anything that is not exact and certain, and (ii) refer to one specific interpretation of the notion. In this document, *uncertainty* would bear the former meaning when used on its own, and adopt the latter interpretation otherwise (i.e., *uncertain* knowledge versus *incomplete* knowledge).

is addressed by allowing a piece of information to be incomplete, thereby permitting a previously derived conclusion to be revised in light of new information.

The use of symbolic modalities and the ordinal representation of uncertainty can be more natural and appealing in certain cases. However by the same token, purely symbolic approaches are not able to produce answers with a precise degree of certainty which is a highly desirable feature in many applications. Furthermore, from a computational point of view, reasoning with a numerical method is usually more efficient than with a symbolic approach, explaining the predominant application of quantitative reasoning in practice. To this end, apart from purely numerical systems such as belief networks [39], valuation networks [3, 53, 54] and evidential networks [64, 67], attempts have also been made to devise systems that tightly integrate logical techniques with numerical representation of uncertainty such as probabilistic logic [36], possibilistic logic [14] and subjective logic [27], or hybrid systems that combine symbolic and numerical approaches in a loosely coupled manner such as ATMS[2] augmented with uncertainty management (e.g., probabilistic ATMS [55], possibilistic ATMS [13], and ATMS using Dempster-Shafer methods [30]), and argumentation systems enhanced with uncertainty treatment (e.g., argumentation systems that incorporate probabilistic information [22] or Dempster-Shafer belief functions [24, 31]).

Since the efficiency of the reasoning process and the ability to distinguish the degrees of certainty among hypotheses are critical to effective decision making, numerical methods are adopted as the representation and reasoning basis to tackle uncertainty in Influx$_1$. In order to further explore the types of uncertainty and associated formalisms that might be suited to Influx$_1$, we have considered the three major theories which are at the foundation of the aforementioned reasoning systems and others: probability theory, Dempster-Shafer theory and possibility theory. A brief introduction to the theories is given below.

2.1 Probability theory

In this mathematical theory, the uncertainty of information is defined by means of probabilities. To date, the notion of probability has admitted two widely accepted interpretations, namely, *objective* and *subjective* probability. From the viewpoint of the frequentists, probability is concerned with chance and randomness (and thus, objective). A typical example of an objective probability is the relative frequency of a particular outcome of an experiment, provided that there is a sufficient (or infinite) number of outcomes that can be observed. Conversely, from the standpoint of the subjectivists, probability can be associated with non-repeatable events; in which case a probability is supposed to reflect the subjective belief of an agent for the problem at hand (based on its experience and/or current state of information).

As preliminaries, let S denote the *sample space* which is is the set of all possible outcomes associated with an experiment (e.g., the set of possible outcomes for tossing two coins is $S = \{HH, HT, TH, TT\}$). Often, one is interested in a certain aspect of the experiment's outcome (such as the number of heads), and thus may wish to represent the sample space accordingly. This can be done via the use of a random variable $X : S \to \Theta$ where Θ

[2] Assumption-based Truth Maintenance System

is a measurable space (such as $\Theta = \{0, 1, 2\}$, the possible number of heads in such an experiment). The space Θ is also referred to as the *state space* of X. Generally, one can define multiple state spaces corresponding to a sample space, and a sample space in one context can be a state space in another context and vice versa.

Probability theory measures the uncertainty of an experiment by assigning a probability $P(A)$ to each event $A \subseteq S$. These probabilities must obey the following three basic axioms (known as the *Kolmogorov* axioms for probability):

$$P(A) \geq 0 \tag{1}$$

$$P(S) = 1 \tag{2}$$

$$\text{if } A \cap B = \emptyset \text{ then } P(A \cup B) = P(A) + P(B). \tag{3}$$

A function that assigns a probability to each of the elements of a sample or state space (e.g., for unbiased coins: $P(HH) = P(HT) = P(TH) = P(TT) = \frac{1}{4}$) is referred to as a *probability distribution*.

Since any event A can be described as the union of the joint events $(A,\ B_i)$, the third axiom also allows the probability for A to be computed from the probabilities associated with the joint events:

$$P(A) = \sum_{i=1}^{n} P(A, B_i) \tag{4}$$

where B_i, $i = 1,\ 2, \ldots,\ n$, constitute a set of exhaustive and mutually exclusive events. A probability distribution defined on such joint events $(A_i,\ B_j)$ is known as a *joint probability distribution*. For example, given the joint probability distribution in Table 1, the probability associated with the occurrence of '$virus_A$' is computed as:

$$P(virus_A) = \textstyle\sum_{i=1}^{n} P(virus_A,\ incident_i) = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}.$$

	$virus_A$	$virus_B$
$incident_1$	$\frac{1}{2}$	$\frac{1}{9}$
$incident_2$	$\frac{1}{6}$	$\frac{2}{9}$

Table 1: An example joint probability distribution.

A probability distribution defined on a set of events is often not static and can be updated in the light of new evidence. As evidence in probability theory is often an observation showing that a specific event has occurred, a standard way to perform updating is by making use of conditional probability functions. A conditional probability $P(A \,|\, B)$ is a function which specifies the probability that A will occur, given the condition that B is known to have occurred and *everything else is irrelevant for* A. In probability theory, conditional probabilities are defined in terms of joint events using *Bayes' rule of conditioning*:

$$P(A \mid B) = \frac{P(A\,,\,B)}{P(B)}. \tag{5}$$

Suppose that $incident_1$ in Table 1 is observed, the probability associated with the occurrence of $virus_A$ now becomes

$$P(virus_A \mid incident_1) = \frac{P(virus_A\,,\,incident_1)}{P(incident_1)} = \frac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{9}} = \frac{9}{11}. \tag{6}$$

Here, if $P(B \mid A)$ $=P(B)$, A and B are said to be independent. Similarly, if $P(B \mid C, A)$= $P(B \mid C)$, A and B are said to be conditionally independent given C. Closely related to Bayes' rule of conditioning (but with a different emphasis) is the well-known *Bayes' theorem* defined as

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)} \tag{7}$$

According to the above formula, the probability associated with the occurrence of $incident_1$ given $virus_A$ can be computed as follows:

$$P(incident_1 \mid virus_A) = \frac{P(virus_A \mid incident_1)\,P(incident_1)}{P(virus_A)} = \frac{\frac{9}{11}(\frac{1}{2}+\frac{1}{9})}{\frac{1}{2}+\frac{1}{6}} = \frac{3}{4}. \tag{8}$$

Bayes' theorem is widely used both in theory and applications. The theorem is particularly useful when the probabilities associated with joint events are absent (which is often the case in practice) but conditional probabilities (e.g., $P(A \mid B)$ and $P(B \mid A)$) and prior probabilities (e.g., $P(A)$ and $P(B)$) are available. Those models which utilise Bayes' theorem in their computation are commonly referred to as *Bayesian models.*

Though probability theory is a well-recognised and widely-used mathematical tool to handle uncertainty, the theory has a number of limitations. Our major concerns in utilising probability theory as a tool to handle uncertainty in Influx_1 include the following issues:

- A probabilistic model requires the existence of a probability distribution to capture the epistemic state of an agent[3]. When such a probability distribution is unavailable (thus, the agent is completely ignorant about the situation), it is a standard practice to apply the *principle of insufficient reason* (corresponding to the maximum entropy principle) which effectively assigns a uniform distribution to the state space Θ. This can cause some difficulties in interpretation and manipulation of the defined probabilities. For instance, a probabilistic model may not be able to distinguish between whether the uncertainty is due to the lack of information (incomplete information) or to the variability of past results [14] (as illustrated in Example 1), or may be susceptible to epistemic inconsistency when the granularity of Θ is manipulated (as

[3] An agent can be any entity that has the capability to perceive and reason about the current state of the world.

shown in Example 2). Examples 1 and 2 presented below re-illustrate historical arguments in the relevant context.

> **Example 1**: Let $\Theta_1 = \{virus_A, virus_B\}$ denote the set of all types of virus possibly responsible for a security incident. Given the uniform probability distribution $P(virus_A) = P(virus_B) = \frac{1}{2}$, one would have difficulty in determining whether this information is (a) a statement of total ignorance, or (b) the result of accumulating an equal amount of evidence supporting both types of virus, respectively.
>
> **Example 2**: Suppose that a security analyst distinguishes between the two subtypes of $virus_A$ (i.e., $virus_A$ now corresponding to $\{virus_{A1}, virus_{A2}\}$), hence $\Theta_2 = \{virus_{A1}, virus_{A2}, virus_B\}$. Applying the principle of insufficient reason on Θ_2 would yield
>
> $$P(virus_{A1}) = P(virus_{A2}) = P(virus_B) = \frac{1}{3}$$
>
> which is, in the case of total ignorance, incompatible with those in Example 1. In the same vein, imagine that the analyst would like to discern between $virus_A$ and $virus_B$ only, and thus mapping Θ_2 into Θ_1. Using the standard coarsening algorithm in [46], the analyst now obtains $P(virus_A) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}$ and $P(virus_B) = \frac{1}{3}$. To this end, it appears that the support for $virus_A$ is generated by itself in the absence of any evidence and known prior probabilities.

- A probabilistic method is also not able to distinguish between *disbelief* and *lack of information*, as shown in the following example, due to the unity probability distribution imposed on a proposition and its negation.

> **Example 3**: In another scenario, the analyst is interested in identifying if a security incident is caused by $virus_A$, thus $\Theta_4 = \{virus_A, \overline{virus_A}\}$. Being able to recognise only some features that may suggest the involvement of $virus_A$ and *nothing else*, the analyst specifies $P(virus_A) = \frac{1}{3}$. Under a probabilistic interpretation, the probability for the negation of the proposition can be automatically inferred: $P(\overline{virus_A})$ = 1 - $P(virus_A) = \frac{2}{3}$, which does not faithfully reflect the belief of the analyst in this particular situation (since the rest of the features may or may not suggest the possible involvement of $virus_A$).

The limitations presented above are among the main reasons for alternative theories to be proposed and developed. Representatives among them are Dempster-Shafer theory and possibility theory which will now be introduced.

2.2 Dempster-Shafer (D-S) theory

The so-called Dempster-Shafer (D-S) theory[4] originates from the mathematical framework concerning *lower and upper probabilities* by Dempster [5] which extended classical prob-

[4]The Dempster-Shafer theory is also referred to as the *theory of belief functions*.

ability theory, subsequently formulated by Shafer as *the mathematical theory of evidence* [46], and further studied by Smets in the framework of *transferable belief models* [51].

Comparisons of D-S theory and its predecessor Bayesian probability theory have contributed to many debates in the literature of evidential reasoning; with a consensus acknowledging the merits of both (c.f., [50]), thus any assessment of their adequacy and suitability should be subject to the specific scenarios and the problems at hand. In this regard, most relevant to our concerns is that D-S theory, unlike the Bayesian approach, tolerates the *incompleteness* of information. In other words, the theory does not require a complete probabilistic model to be specified, and thus is more suitable when such a model is only partially available. To this end, the theory preserves the numerical representation for uncertainty (in agreement with probability theory), but allows for the modelling of incompleteness by means of disjunctive sets (just as in classical logics). Consequently, instead of tackling the elements in the state space Θ themselves (as in the Bayesian formalism), D-S theory deals with all possible subsets of Θ (i.e., elements of 2^{Θ}). The discussion that follows deals primarily with the applied aspect of the theory.

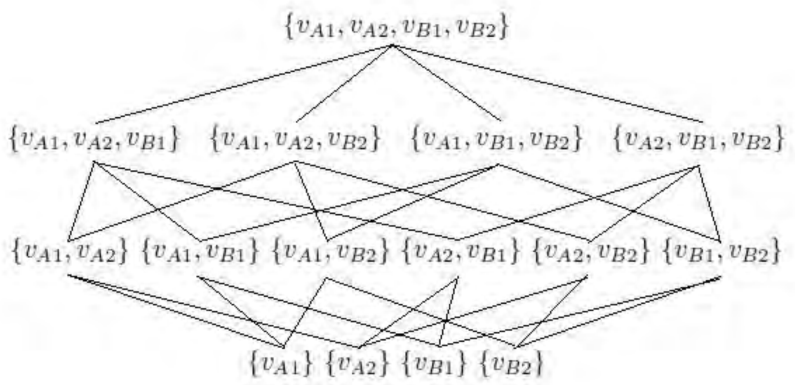


Figure 2: A discernment space containing all the possible propositions derived from $\Theta = \{v_{A1}, v_{A2}, v_{B1}, v_{B2}\}$.

Formally, let a *frame of discernment* Θ be a set of possible answers to some question, or possible values for some variable[5]. The elements of Θ are required to be *exhaustive* and *mutually exclusive*; that is, at any one time, one and only one element in Θ can be true. A so-called *proposition* or *hypothesis*, A, is defined as a subset of Θ, and is said to be true if the truth lies within A. For example, let A correspond to the proposition 'the incident is caused by a virus of type A' (see Table 1). Then A would be represented as $A = \{virus_{A1}, virus_{A2}\}$, and would be true if either $\{virus_{A1}\}$ or $\{virus_{A2}\}$ is true. The set of all possible propositions derived from Θ constitutes a *discernment space* and is denoted as 2^{Θ}.

Figure 2 illustrates the space of discernment derived from a given frame of discernment Θ. Among the depicted propositions, the root proposition is Θ itself, and those at the bottom

[5]A frame of discernment in D-S theory corresponds to a state space in probability theory.

are called *singletons* (i.e., one-element propositions). Please note that the proposition corresponding to the *empty set*, $\emptyset$, is not included in Figure 2.

The primitive function in D-S theory is a function $m : 2^\Theta \to [0, 1]$, called a *basic probability assignment* (bpa)[6], such that:

$$\begin{gathered} \textstyle\sum_{A \subseteq 2^\Theta} m(A) = 1, \\ m(\emptyset) = 0 \end{gathered} \tag{9}$$

where $m(A)$ encodes the belief mass assigned to the proposition A.[7]

A bpa can be considered a generalisation of the probability function in that one can assign belief mass to elements of the space of discernment, 2^Θ, rather than elements of the frame of discernment Θ, as in the Bayesian fashion. However, it is important to note that one should *not* simply interpret $m(A)$ as the probability of the occurrence of A, as $m(A)$ is the belief mass assigned to A *only* and to none of its subsets[8] (see Example 4).

> **Example 4**: Given $\Theta = \{virus_{A1}, virus_{A2}, virus_B\}$ which denotes all the possible causes for an incident I, and a bpa m reported by an agent as follows
>
> $$\begin{gathered} m(\{virus_{A1}\}) = \tfrac{1}{2}, \\ m(\{virus_{A1}, virus_{A2}\}) = \tfrac{1}{3}, \\ m(\Theta) = m(\{virus_{A1}, virus_{A2}, virus_B\}) = \tfrac{1}{6}, \end{gathered}$$
>
> and $m(A) = 0$ for all other $A \subset \Theta$ (note that $\sum m(A) = 1$).
> The quantity $\frac{1}{2}$ should be interpreted as the portion of belief *committed* to $\{virus_{A1}\}$, while the quantity $\frac{1}{3}$ is the portion of belief *pending* over $\{virus_{A1}, virus_{A2}\}$ (which is not yet committed to either $\{virus_{A1}\}$ or $\{virus_{A2}\}$ due to lack of knowledge, but can be gradually transferred to $\{virus_{A1}\}$ or $\{virus_{A2}\}$ in light of further evidence). Likewise, the quantity $1 - \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ is the portion of belief pending over the whole frame (which remains unassigned after allocation of belief to various proper subsets of Θ). When $m(\Theta) = 1$ (i.e., when there is no belief mass assigned to any of the proper subsets of Θ), the agent is said to be completely ignorant. In such a case, the agent believes that the incident I is *certainly* caused by one of the elements of Θ, but it does not have any knowledge to make a more informed judgment (thus, $m(\Theta) = 1$).

With such expressiveness, D-S theory does not experience the aforementioned difficulties encountered by the Bayesian formalisms due to the fact that incompleteness of information can be expressed explicitly by assigning belief mass to unions of singletons or to the whole frame of discernment Θ. More specifically, by representing the state of total ignorance

[6] The basic probability assignment m is also widely referred to, in its unnormalised form, as *basic belief assignment* (*bba*) by Smets [49], or more generally, a *mass assignment*.

[7] Shafer adopts a closed-world assumption and enforces $m(\emptyset)$ to be 0. In the open-world formulation of D-S theory studied by Smets [51], belief mass can be assigned to $\emptyset$ to indicate the fact that the answer for a problem may lie outside those captured in the current frame of discernment.

[8] For this reason, a bpa is not required to be inclusion-monotonic [14]: it is plausible to have $m(\{virus_{A1}\}) > m(\{virus_{A1}, virus_{A2}\})$ even though $\{virus_{A1}\} \subset \{virus_{A1}, virus_{A2}\}$.

as $m(\Theta) = 1$, D-S theory maintains the difference between the variability of past results ($m(\{virus_{A1}\}) = m(\{virus_{A2}\}) = m(\{virus_B\}) = \frac{1}{3}$) and incompleteness ($m(\Theta) = 1$) discussed in Example 1. As $m(\Theta) = 1$ resists changes in the granularity of Θ, the epistemic inconsistency illustrated in Example 2 is also eliminated (transformation between Θ_2 and Θ_1 does not alter the state of ignorance of the discussed bpas). Likewise in Example 3, the theory is able to clearly distinguish between the states of disbelief and lack of information. In this case, since the analyst does not have any evidence that *disconfirms* $\{virus_A\}$ (i.e., confirms $\{\overline{virus_A}\}$), this epistemic state can be faithfully reflected in a D-S model by allocating the rest of the total belief (1 - $\frac{1}{3}$) to $m(\Theta)$ (instead of $\{\overline{virus_A}\}$).

A bpa induces two other important set-functions, respectively known as a belief function *Bel* and a plausibility function *Pl*. *Bel* and *Pl* are defined by:

$$Bel(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B), \tag{10}$$

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B). \tag{11}$$

The D-S belief function, *Bel*, corresponds to a probability function in probability theory. *Bel* also obeys basic axioms for probabilities (see Section 2.1) except that it is not additive (i.e., Eq. (3) is removed):

$$Bel(A) \geq 0,$$
$$Bel(\Theta) = 1,$$
$$\text{if } A \cap B = \emptyset \text{ then } Bel(A \cup B) \geq Bel(A) + Bel(B).$$

Dismissing the additive axiom when dealing with beliefs is a major difference between D-S and Bayesian models. For example, one can calculate the *Bel* and *Pl* values for various propositions in Example 4 using Eq. (10) and Eq. (11) as

$$Bel(\{virus_{A1}\}) = \frac{1}{2},$$
$$Pl(\{virus_{A1}\}) = \frac{1}{2} + \frac{1}{3} + \frac{1}{6} = 1,$$
$$Bel(\{virus_{A1}, virus_{A2}\}) = \frac{1}{3} + \frac{1}{2} = \frac{5}{6},$$
$$Pl(\{virus_{A1}, virus_{A2}\}) = \frac{1}{3} + \frac{1}{6} + \frac{1}{6} = 1,$$
$$Bel(\{virus_B\}) = 0,$$

and

$$Pl(\{virus_B\}) = \frac{1}{6}.$$

It is easy to see from the above example that both the belief and plausibility functions are non-additive (e.g., $Bel(\{virus_{A1}, virus_{A2}\}) \geq Bel(\{virus_{A1}\}) + Bel(\{virus_{A2}\})$, and that $Bel(\Theta)$ and $Pl(\Theta)$ are always equal to 1.

In essence, Bel measures the extent to which an agent *believes* A, and Pl measures the degree to which the agent *fails to disbelieve* A (i.e., A *could be true* and $Pl(A)$ = 1 - $Bel(\overline{A})$). Subsequently, the size of the interval [$Bel(A)$, $Pl(A)$] can be regarded as the amount of uncertainty with respect to A, given the evidence. It is belief that is committed by the evidence to neither A nor $\overline{A}$. The larger the interval is, the more likely the information or evidence is missing or unreliable. When an agent is in a state of total ignorance (i.e., $m(\Theta) = 1$, and thus $Bel(A) = 0$ and $Pl(A) = 1$), the interval [$Bel(A)$, $Pl(A)$] would be [0, 1]. At the other extreme, if the agent has complete knowledge about the situation for it to allocate mass to the singletons of Θ only, the probability, belief and plausibility measures would become identical ($P(B) = Bel(B) = Pl(B)$ for any $A \subseteq \Theta$) as well as additive ($Bel(A_1 \cup A_2) = Bel(A_1) + Bel(A_2)$ if $A_1 \cap A_2 = \emptyset$). In this case, the interval [$Bel(A)$, $Pl(A)$] would collapse into $P(A)$, and reasoning with D-S belief functions reduces to standard Bayesian reasoning.

Like Bayesian models, beliefs in D-S models can be updated via conditioning where the notion of conditional belief is defined as:

$$Bel(A \mid B) = \frac{Bel(A \cup \overline{B}) - Bel(\overline{B})}{1 - Bel(\overline{B})}$$

and

$$Pl(A \mid B) = \frac{Pl(A \cap B)}{Pl(B)}$$

$Bel(A \mid B)$ and $Pl(A \mid B)$ are collectively referred to as Dempster's rule of conditioning which effectively reduces to $P(A \mid B)$ when probability functions are in place of belief functions.

Whilst rectifying limitations of Bayesian probability theory, D-S theory is itself not free of issues. Most notoriously, the expressiveness offered by a D-S model obviously comes at the cost of higher computational complexity. Limitations of D-S models are clarified, discussed and addressed in Section 3 of this document.

2.3 Possibility theory

The formal theory of possibility was developed based on the notion of fuzzy sets by Zadeh [68], though the idea of using possibility measures, as an alternative to probabilities, is generally attributed to Shackle's work [45] on modelling expectation and potential surprise in human decision making. Possibility theory addresses uncertainty from another perspective. Instead of viewing uncertainty from a statistical standpoint as in probability theory, possibility theory focuses on the uncertainty intrinsic in linguistic information. That is, the theory is mainly concerned with the meaning of the information (rather than its measure) [68] and tackles problems such as ambiguity and vagueness.

Taking place of a probability distribution, a possibility distribution is a function π: $2^\Theta \rightarrow [0, 1]$ such that $\pi(\emptyset) = 0$ and $\pi(\Theta) = 1$. The quantity $\pi(A) = 1$ means that A is totally possible (plausible), while $\pi(A) = 0$ implies that A is impossible. In the absence of information, possibility theory adopts the *principle of minimum specificity* which assigns to each state the highest degree of possibility. Table 2 illustrates an example possibility distribution in contrast with a probability distribution.

The plausibility and certainty of a proposition A (in terms of distance to an ideally plausible situation) can be evaluated using the *possibility measure* Π and *necessity measure* N:

$$\Pi(A) = \max_{s \in A} \pi(s),$$

$$N(A) = 1 - \Pi(\bar{A}) = \min_{s \notin A}(1 - \pi(s));$$

and

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)),$$

$$N(A \cap B) = \min(N(A), N(B)).$$

Though focusing on a different aspect of uncertainty, measures of uncertainty in possibility theory hold some connections with those in probability and D-S theories. Whereas a probability distribution provides precise and disjunct pieces of information, a possibility distribution encodes imprecise, but consonant[9], pieces of information [14]. As a result, if A is impossible, it is likely to be improbable, but a high degree of possibility does not entail a high degree of probability, nor does a low degree of probability indicate a low degree of possibility [68]. For example, as illustrated in Table 2, while it is totally possible that the incident is caused by either $\{virusA_1\}$ or $\{virusA_2\}$, it is much more likely that the incident is caused by $\{virusA_1\}$ rather than $\{virusA_2\}$. In general, their relationship can be captured as $N(A) < P(A) < \Pi(A)$, $\forall A$.[10] With regard to D-S theory, both theories

x	$\{virusA_1\}$	$\{virusA_2\}$	$\{virusB\}$
$\pi(x)$	1	1	$\frac{2}{3}$
$P(x)$	$\frac{3}{4}$	$\frac{3}{16}$	$\frac{1}{16}$

Table 2: A possibility distribution π in comparison to a probability distribution P.

correspond to a particular family of probability measures [14]. In particular, the possibility and necessity measures coincide with the plausibility and belief functions, with $Pl \geq \Pi$ and $Bel \leq N$ (in the general case) and $Pl = \Pi$ and $Bel = N$ (in the finite consonant case) [14].

Possibility theory provides a simple but useful approach to handle the uncertainty and incompleteness of information. Nevertheless, our focus in designing and developing Influx$_1$ is on modelling the *degree of belief* (or *confidence*) regarding the truth of a piece of information, rather than *degree of truth* associated with vague and ambiguous information. Therefore although ideas and models based on possibility theory can be potentially integrated into various elements of the reasoning tool in the future, the theory is not considered as a theoretical basis for the current design and implementation of Influx$_1$.

[9] Pieces of information are referred to as *consonant* if they can be represented in a nested structure.

[10] Nevertheless, transformation between a probability and possibility measure can be done in certain circumstances, generally with some information loss (c.f., [14]).

2.4 Uncertainty formalism adopted in Influx

Influx is intended to be a reasoning system dealing with *imperfect* information (and knowledge). Of particular interest in regard to the anticipated imperfection are *incompleteness* (and thus, *imprecision*) (i.e., a piece of information is not sufficiently specific for an agent to answer a relevant question), *uncertainty* (i.e., a piece of information is not known to be true or false) and *contradiction* (i.e., pieces of information provided from different sources can be contradicting). Thus, unlike the taxonomy presented in Figure 1, uncertainty is not considered a subtype of incompleteness in Influx. Instead, uncertainty and incompleteness, together with contradiction, characterise a more general notion of imperfectness. To represent and reason with such imperfect information, a numerical formalism (specifically, D-S theory) is adopted as the theoretical framework for the development of reasoning capabilities in Influx.

Though the Bayesian formalism is convenient and efficient in tackling uncertainty, D-S theory is chosen due to a number of factors. D-S theory provides a more flexible reasoning framework, allowing one to represent forms of uncertainty difficult to realise within the framework of probability theory (such as the incompleteness/imprecision of information, besides the uncertainty). In addition, the literature of evidential reasoning provides a large collection of DS-related combination rules to fuse information from multiple sources most appropriately for the situations at hand (as will be discussed in Section 3). Indeed, with an ability to represent all forms of uncertainty, from total ignorance to full knowledge, a DS-based reasoning tool can achieve high versatility — being able to accommodate a wider range of applications and catering for various reasoning scenarios. For instance, D-S theory, in principle, allows for reasoning using belief functions in situations where only partial knowledge is available, seamlessly reducing into Bayesian reasoning when full knowledge is available, and possibly collapsing into standard rule-based inferencing when such a form of reasoning is desired. Ultimately, each of these reasoning formalisms (and possibly others) can prevail in different conditions and scenarios — it is highly desirable to have at one's disposal a unified framework capable of supporting multiple reasoning modalities including those presented above. Our decision to adopt D-S theory as a basis for representing and reasoning with uncertainty is driven by this ultimate vision.

It is important to note that throughout the design and development of Influx, various restrictions on the expressiveness of the D-S calculus have been, and are to be, made in order to enhance the practicability of the reasoning tool. In this aspect, D-S theory offers one the freedom to flexibly derive and apply restrictions that best suit specific goals and scenarios.

3 Preliminaries of the Dempster-Shafer theory

This section is devoted to a brief discussion of the applied aspect of D-S theory in the context of information fusion. Recall that given a finite frame of discernment Θ which captures all the possible and mutually exclusive answers for a problem at hand, a bpa m is a primitive function that allows an agent to express its present state of belief by assigning belief mass to various propositions A such that

$$m : 2^{\Theta} \to [0,1],$$
$$\sum_{A \subseteq 2^{\Theta}} m(A) = 1, \text{and}$$
$$m(\emptyset) = 0.$$

Generally, not every subset A of Θ would receive a non-zero belief assignment. When $A \subseteq \Theta$ and $m(A) > 0$, A is referred to as a *focal set*. Bpas may be described according to their focal sets. More specifically, a bpa m is said to be:

- *vacuous*, if $m(\Theta) = 1$ (which represents the agent's state of total ignorance);
- *dogmatic*, if Θ is not a focal set;
- *simple*, if it has at most two focal sets, including Θ;
- *dichotomous*, if it has at most three focal sets (A, $\bar{A}$ and Θ);
- *categorical*, if it is simple and dogmatic;
- *normal*, if $\emptyset$ is not a focal set, and subnormal otherwise;
- *consonant*, if all its focal sets $A_1, ..., A_N$ are nested: $\emptyset \subseteq A_1 \subseteq ... \subseteq A_N \subseteq \Theta$; and
- *Bayesian*, if all of its focal sets are singletons: $m(A) > 0$ and $|A| = 1$.

When there is more than one bpa defined on the same frame of discernment, their effects need to be combined. Shafer [46] in developing the theory has reformulated Dempster's rule of combination (now commonly referred to as the Dempster-Shafer, or D-S rule) to combine evidence, provided that the bpas being combined are induced by *distinct* and *independent* pieces of evidence.

In many real-world applications dealing with uncertainty, the reliability of information is often enhanced by means of collecting relevant pieces of information, often with varying degrees of certainty, from multiple sources, necessitating the availability of methods to aggregate the beliefs, each associated with each piece of information. In this aspect, the D-S rule has turned out to be very useful, and thus has been widely investigated in the field of information fusion to address multiple facets of uncertainty. Using the D-S rule to combine an agent's present state of belief (expressed in m_1) with a specific state of belief induced by a new piece of evidence (expressed in m_2) allows the agent's state of belief to be updated in light of the new evidence. Similarly, fusing a number of 'uncertain' opinions from independent experts is likely to provide a more trustworthy answer for a problem of interest, and combining various 'incomplete' pieces of knowledge would potentially yield a more 'complete' piece of knowledge about a situation.

In the framework of D-S theory, combining independent beliefs pertaining to the same piece of information is usually carried out through an aggregation of the corresponding bpas that describe them. Let m_1 and m_2 be the two bpas to be combined, the combined bpa $m_1 \otimes m_2$ (or m_{12}) is computed using the D-S rule as follows

$$m_{12}(\emptyset) = 0 \text{ and } \forall A \in 2^{\Theta} \backslash \emptyset :$$

$$m_{12}(A) = \frac{1}{1 - K_{12}} \sum_{B,C \in 2^{\Theta}, B \cap C = A} m_1(B) m_2(C). \tag{14}$$

where the conflict mass K is:

$$K_{12} = m_{12}(\emptyset) = \sum_{B,C \in 2^{\Theta}, B \cap C = \emptyset} m_1(B) m_2(C)$$

In the above formula, K represents a basic probability mass associated with conflicts among the sources of evidence (which are m_1 and m_2 in this example). If K_{12} is close to zero, the bpas are not in conflict, whereas if K_{12} is close to 1, the bpas are in conflict. When $K_{12} = 1$, the bpas are said to be in total conflict and thus *non-combinable*. Eq. (14) is also referred to as an 'orthogonal sum' due to the way it combines belief masses provided by the evidence. Example 5 presents a simple application of the D-S rule to information fusion, while Figure 3 graphically illustrates the combination process.

Example 5: Let $\Theta = \{virus_{A1},\ virus_{A2},\ virus_B\}$, and the bpas m_1, m_2 provided by the two independent sensors 1 and 2 be as follows:

$$m_1(\{virus_{A1}\})\text{=}0.3,\ m_1(\{virus_B\})\text{=}0.5,\ m_1(\Theta)\text{=}0.2, \text{ and}$$
$$m_2(\{virus_{A1},\ virus_{A2}\})\text{=}0.7,\ m_2(\Theta)\text{=}0.3.$$

In this example, $\{virus_{A1}\}$ and $\{virus_B\}$ are focal elements of m_1 while $\{virus_{A1},\ virus_{A2}\}$ is the only focal element of m_2. The information provided by sensors 1 and 2 are both uncertain and incomplete (the type of information that a Bayesian formalism would find difficult to represent in an adequate manner). In particular, the information provided by sensor 2 illustrates a typical scenario where the sensor can predict that the virus is much more likely to be of type A (rather than type B), but is incapable of detecting the specific virus. Bringing together the effect of m_1 and m_2 on Θ (see Figure 3), the combined bpa m_{12} produced by an application of the D-S rule clearly indicates a support for $\{virus_{A1}\}$ as the possible answer:

$$\begin{aligned} m_{12}(\{virus_{A1}\}) &= 0.46, \\ m_{12}(\{virus_B\}) &= 0.23, \\ m_{12}(\{virus_{A1}, virus_{A2}\}) &= 0.22, \\ m_{12}(\Theta) &= 0.09; \end{aligned}$$

The D-S rule also possesses attractive algebraic properties: it is both *associative* ($m_1 \otimes (m_2 \otimes m_3) = (m_1 \otimes m_2) \otimes m_3)$) and *commutative* ($m_1 \otimes m_2 = m_2 \otimes m_1$). This means that the combination process for a large number of bpas is not required to be computed at once (batch), but can be conducted in a sequential, incremental and local manner.

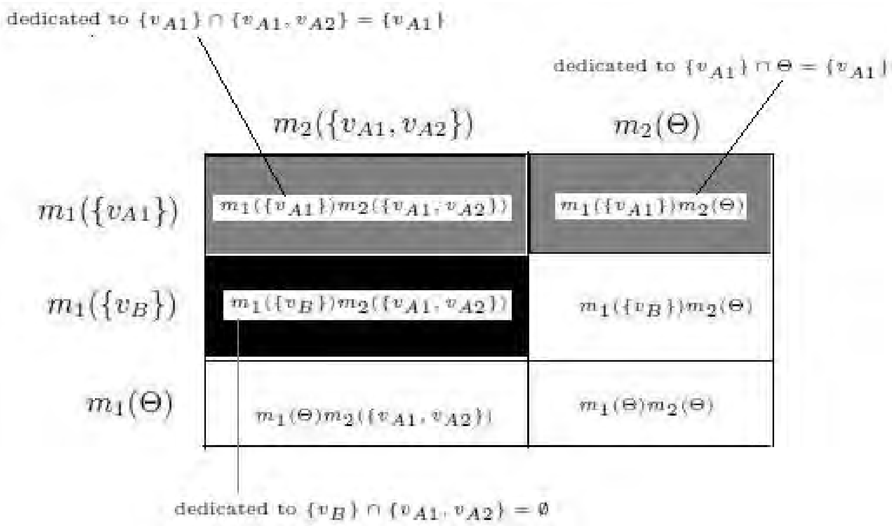


Figure 3: A graphical representation of the orthogonal sum in Example 5 where $\{v_{A1}\}$, $\{v_{A2}\}$ and $\{v_B\}$ correspond to $\{virus_{A1}\}$, $\{virus_{A2}\}$ and $\{virus_B\}$, respectively.

Frame coarsenings and refinements In situations where the elements in a frame of discernment Θ are not specific enough to deal with the problem at hand, Θ can be transformed into a more refined frame Ω using the following mapping rules:

$$\omega : 2^{\Theta} \rightarrow 2^{\Omega},$$

$$\omega(A) = \bigcup_{\theta \in A} \omega(\{\theta\}) \tag{15}$$

where the sets $w(\{\theta\})$ constitute a disjoint partition of Ω. Shafer [46] called such a mapping ω a *refining*, Ω a *refinement* of Θ, and Θ a *coarsening* of Ω. In a similar fashion, it is possible to derive a finer or coarser bpa by discerning a bpa on Ω or Θ ($\omega : 2^{\Theta} \rightarrow 2^{\Omega}$) using the following formulas:

$$m_{\Omega}(\omega(A)) = m_{\Theta}(A), \forall A \subset \Theta,$$

$$m_{\Theta}(A) = \sum_{\substack{B \subset \Omega, \\ A = \omega(B)}} m(B) \tag{16}$$

Example 6: Given the frame of discernment $\Theta = \{virus_{A1}, virus_{A2}, virus_B\}$ and the following bpa expressed on Θ:

$$\begin{aligned} m_\Theta(\{virus_{A1}\}) &= \frac{1}{3}, \\ m_\Theta(\{virus_{A2}\}) &= \frac{1}{3}, \\ m_\Theta(\{virus_B\}) &= \frac{1}{3}, \end{aligned}$$

let us define the two following refinings ϕ and ω as

$$\begin{aligned} \omega &: 2^\Theta \to 2^\Omega, \\ \omega(\{virus_{A1}\}) &= \{virus_{A1}\}, \\ \omega(\{virus_{A2}\}) &= \{virus_{A2}\}, \\ \omega(\{virus_B\}) &= \{virus_{B1}, virus_{B2}\}, \end{aligned}$$

$$\begin{aligned} \phi &: 2^\Phi \to 2^\Theta, \\ \phi(\{virus_A\}) &= \{virus_{A1}, virus_{A2}\}. \\ \phi(\{virus_B\}) &= \{virus_B\}. \end{aligned}$$

The refinement of Θ resulting from the application of the refining ω to Θ is

$$\Omega = \{virus_{A1}, virus_{A2}, virus_{B1}, virus_{B2}\},$$

and the coarsening of Θ according to the refining ϕ is

$$\Phi = \{virus_A, virus_B\}.$$

Likewise, the bpa m_Θ can be discerned on Φ and Ω (using Eq. (15)) as follows

$$\begin{aligned} m_\Phi(\{virus_A\}) &= \frac{2}{3}, \\ m_\Phi(\{virus_B\}) &= \frac{1}{3}; \\ m_\Omega(\{virus_{A1}\}) &= \frac{1}{3}, \\ m_\Omega(\{virus_{A2}\}) &= \frac{1}{3}, \\ m_\Omega(\{virus_{B1}, virus_{B2}\}) &= \frac{1}{3}. \end{aligned}$$

In general, a bpa associated with a frame of discernment is not a Bayesian bpa as represented in the above example, and thus performing frame coarsening or refinement can inevitably result in a certain degree of information loss.

Despite its popularity and appealing features, the D-S rule has received a number of concerns and criticisms regarding its practical usage. A compilation of such concerns include:

- the produced results may be counter-intuitive in the case of very highly conflicting evidence,

- the potentially high computational complexity associated with the D-S rule,
- the bodies of evidence involved in the combination are required to be distinct and independent, and
- all the sources of evidence (e.g., sensors) are assumed to be reliable.

The sections that follow discuss and address these concerns.

3.1 Combination of highly conflicting evidence

A well-known criticism of the D-S rule is by Zadeh [69, 70, 71] who states that the rule yields counter-intuitive results when dealing with highly conflicting pieces of evidence. For instance, fusing the two bpas m_1 and m_2 (given by two independent experts) using the D-S rule, one would obtain the combined bpa m_{12} as illustrated in Table 3. According to

	m_1	m_2	m_{12}
$\{virus_{A1}\}$	0.9	0.0	0.0
$\{virus_{A2}\}$	0.0	0.9	0.0
$\{virus_B\}$	0.1	0.1	1.0
Θ	0.0	0.0	0.0

Table 3: An example combination of highly conflicting dogmatic bpas using the D-S rule.

Zadeh, this fusion result is counter-intuitive because it provides strong support for $virus_B$ which is considered the least likely cause by both the experts.

Mathematically, this seemingly counter-intuitive result is directly caused by the combination of the assignment of zero values to $m_1\{virus_{A2}\}$ and $m_2\{virus_{A1}\}$ (which in effect completely eliminates $\{virus_{A2}\}$ and $\{virus_{A1}\}$ as possible answers) and the use of Bayesian belief functions. Thus, a more careful examination of the illustrated example would reveal that this is more a problem pertaining to general probabilistic analysis, rendering itself a fairly *unfair* judgement of the D-S rule [25]. Nevertheless, the problem can be explained and/or addressed along two main directions. On the one hand, if one looks for consensus among evidence, this result is indeed not counter-intuitive (i.e., the result indicates that $virusB$ is the hypothesis that both sources of evidence agree upon). On the other hand, such counter-intuitive results (if indeed they are) would be resolved if one does not assume the full reliability of the experts' knowledge. For instance, if one assigns a very small value ϵ (say, 0.01) to Θ, the result given by the D-S rule is no longer counter-intuitive as illustrated in Table 4 (a deeper treatment of this issue can be consulted at [25, 4]).

Such an ϵ value is utilised in the implementation of Influx$_1$ in order to avoid having unexpected results in cases where the aforementioned situation arises. Since it is plausible in practice to assume that the belief for a proposition provided by an expert is often

	m_1	m_2	m_{12}
$\{virusA_1\}$	0.9	0.0	0.32
$\{virusA_2\}$	0.0	0.9	0.32
$\{virusB\}$	0.09	0.09	0.35
Θ	0.01	0.01	0.01

Table 4: An example combination of highly conflicting non-dogmatic bpas using the D-S rule and ϵ.

not fully reliable, and the measurements given by a sensor are not totally precise, the assignment of ϵ to Θ is justified.

Furthermore, depending on the specific applications and problems at hand, other combination rules can be utilised that offer different strategies to deal with highly conflicting evidence, e.g., allocating conflicting mass to disjunctive propositions when the information sources are assumed to be unreliable (as in disjunctive rules) or simply assigning it to the empty set to indicate the fact that the answer for a problem may lie outside those captured in the current frame of discernment (as in the unnormalised conjunctive rule). To this end, a collection of such combination rules has been implemented in Influx$_1$ which allows users to resolve the problem of highly conflicting information in a most intuitive and appropriate way.

3.2 Computational complexity

As mentioned in Section 2.2, a major criticism of D-S models is their high computational complexity — dealing with subsets of Θ rather than elements of Θ. Therefore, the computational complexity is exponential to the size of the frame of discernment. Combining multiple pieces of evidence expressed on a large frame of discernment is often intractable. Researchers have attempted to tackle this problem along the following dimensions: (i) cutting down the number of focal elements involved in the combination, (ii) reducing the size of the frame of discernment, and (iii) imposing certain restrictions on the discernment space, thereby allowing the devising of combination algorithms with lower computational complexity. Methods in the three mentioned categories are elaborated further in the following sections.

Other approaches also exist that are not concerned with the size of the frame of discernment or the structure of the discernment space, but adopts specific techniques to reduce the time associated with the combination process. Techniques adopted in such approaches include the Monte-Carlo algorithm (the combined beliefs are estimated by means of random sampling of the possible values of the mass functions to be combined) [61, 62], the Fast Möbiũs Transform (fast transform between different types of belief functions is performed, allowing for beliefs being combined in their most convenient form) [29], efficient set representation and operations (the computational complexity of the combination process is reduced due to a novel representation of focal elements using techniques from finite set theory) [37] and resource-bounded schemes (beliefs are combined in a progressive manner so that the achieved accuracy of the result is proportional to the resource (time) available for the combination process) [23, 21].

3.2.1 Reducing the number of focal elements

As the possible number of focal elements in a bpa can be up to 2^n (where n is the size of the frame of discernment), reducing the number of focal elements in the bpas can significantly reduce the amount of computation required. Approaches pertaining to this category include the Bayesian approximation [59], K-l-x [57], summarisation [32], D1 approximation [2], and inner and outer clustering [6] methods. While the Bayesian approximation method reduces a given bpa to a probability distribution, the other methods focus on producing a modified bpa which satisfies pre-defined constraints specified on the mass value of the focal elements. As a consequence, the K-l-x, summarisation, D1 approximation, and inner and outer clustering methods are data-driven (i.e., the outcomes are determined by the actual mass values assigned to a bpa). The Bayesian approximation method, in contrast, allows the resulting focal elements to be partly determined beforehand. An empirical study of many of these approaches can be found in [2].

3.2.2 Reducing the size of the frame of discernment

An alternative approach to reduce the number of focal elements is to partition the frame of discernment into another frame with smaller size using techniques related to frame coarsening as presented above. Since it is common for a bpa expressed on Θ to include focal elements that are not discerned by Ω, coarsening a frame of discernment generally results in some degree of information loss. To this end, multiple attempts have been made to devise algorithms capable of producing a coarsened frame of discernment from a set of bpas with minimal information loss. More details about coarsening a frame of discernment can be found at [66].

3.2.3 Utilising efficient algorithms to combine evidence

Evidence combination algorithms with reduced computational complexity have been investigated. In particular, Barnett [1], Gorden and Shortliffe [20], Shafer and Logan [47] devised algorithms that achieve linear computational complexity, with certain assumptions imposed on the set of bpas involved in the combination. Barnett's algorithm assumes that only evidence for singleton hypotheses in a frame of discernment is collected, while the Gorden-Shortliffe and Shafer-Logan algorithms do not enforce such a restriction, but require the evidence to have a hierarchical structure as illustrated in Figure 4. In all three approaches, a bpa reported by a sensor is required to be in the form of a simple or dichotomous bpa.

Barnett's algorithm Barnett's algorithm combines pieces of evidence collected for the singleton hypotheses in a frame of discernment. In this algorithm, the evidence provided by each sensor is required to be in the form of a simple or dichotomous bpa. For instance, given $\Theta = \{virus_{A1}, virus_{A2}, virus_{B1}, virus_{B2}\}$, and m_1, m_2, m_3 defined on Θ as:

$$m_1(\{virus_{A1}\}) = 0.2, m_1(\Theta) = 0.8$$

$$m_2(\overline{\{virus_{A2}\}}) = 0.6, m_2(\Theta) = 0.4$$

$$m_3(\{virus_{B1}\}) = 0.3, m_3(\overline{\{virus_{B1}\}}) = 0.2, m_3(\Theta) = 0.5$$

(please note that $\overline{\{virus_{A1}\}} = \{virus_{A2}, virus_{B1}, virus_{B2}\}$ and $\overline{\{virus_{B1}\}} = \{virus_{A1}, virus_{A2}, virus_{B2}\}$). Here, m_1 and m_2 are referred to as simple bpas, and m_3 as a dichotomous bpa[11]. Barnett's algorithm includes mathematical formulas for the combination of these simple and dichotomous bpas in linear time. Details of Barnett's algorithm are described in [1].

The Gorden-Shortliffe algorithm Gorden and Shortliffe [20] proposed an approximation algorithm to compute belief values for the propositions defined in a hierarchical space. A major shortcoming of Barnett's algorithm is that it is only capable of combining pieces of evidence bearing on the singleton propositions defined in a frame of discernment. Gorden and Shortliffe aimed to improve on Barnett's algorithm by allowing pieces of evidence associated with both singleton propositions and (some) disjunctive propositions to be involved in the combination. This can be done by having the relevant discernment space be represented in a strict hierarchical structure. For example, a full graphical representation of the discernment space (i.e., 2^{Θ}) of $\Theta = \{v_{A1}, v_{A2}, v_{B1}, v_{B2}\}$ is given in Figure 2. According to Gorden and Shortliffe, not all the subsets in the structure are semantically meaningful in a particular context. As such, those subsets that are not considered semantically meaningful should be pruned from the structure. In the above example, subsets such as $\{v_{A1}, v_{A2}, v_{B1}\}$, $\{v_{A1}, v_{A2}, v_{B2}\}$, $\{v_{A1}, v_{B1}, v_{B2}\}$, $\{v_{A2}, v_{B1}, v_{B2}\}$, $\{v_{A1}, v_{B1}\}$, $\{v_{A1}, v_{B2}\}$, $\{v_{A2}, v_{B1}\}$, and $\{v_{A2}, v_{B2}\}$ can be considered as not carrying meaningful semantics and thus should be discarded, resulting in a tree-form discernment space illustrated in Figure 4.

Figure 4: A hierarchical structure of subsets of Θ where $\{v_A\} = \{v_{A1}, v_{A2}\}$ and $\{v_B\} = \{v_{B1}, v_{B2}\}$.

The Gorden-Shortliffe algorithm is an approximation in the sense that only nodes in the tree may have a non-zero mass value; a non-zero mass value for any subset that is not represented in the tree will be assigned to its smallest superset that exists in the tree. Details of the Gorden-Shortliffe algorithm are given in [20].

The Shafer-Logan algorithm Shafer and Logan have identified a few disadvantages of the Gordon-Shortliffe approximation algorithm. The disadvantages include: (i) the Gordon-Shortliffe algorithm produces a low-quality result when the degree of conflict among the pieces of evidence is high, (ii) the algorithm is not capable of computing the degree of belief for the negation of a proposition, and thus (iii) the algorithm is not capable

[11] In the terminology used by Shafer [46], a simple bpa corresponds to a simple support function, while a dichotomous bpa corresponds to a simple separable support function.

of computing a plausibility value for the proposition.

Rejecting the need for an approximation algorithm, Shafer and Logan proposed an algorithm for the exact computation of belief values for the propositions. Even though the tree of propositions in the Shafer-Logan algorithm is constructed in a manner similar to the method used in the Gorden-Shortliffe algorithm, the propositions for which their algorithm is capable of computing belief values are not limited to those explicitly represented in the tree. Details of the Shafer-Logan algorithm are given in [47].

The algorithms presented above assume a fixed global frame of discernment, and involve two phases of computation (top-down and bottom-up), and thus are not well suited to implementation in Influx_1. However as will be discussed in subsequent sections, various ideas investigated from the mentioned work have been capitalised in Influx_1 to tackle the problem of computational complexity associated with belief combination such as the adoption of simplified forms of bpas, the utilisation of coarsened (specifically, binary) frames of discernment, and the implementation of only the relevant portion of the discernment space. Generalisation of the Shafer-Logan algorithm is also studied in the context of valuation networks [3, 53, 54] and investigated further in the framework of transferable belief model (TBM) in the form of directed evidential networks [64, 67]. Valuation networks and directed evidential networks will be discussed in the document devoted to a description of Influx_2.

3.3 Unreliability of sources of evidence

By default, sources of evidence are treated equally by a combination rule (i.e., the sources are assumed to be equally reliable). This is not a valid assumption in practical applications, especially those for which Influx_1 is intended. Some examples of this include:

- evidence provided by local/host sensors is more reliable than that provided by remote/network sensors;
- sensors installed at different locations have different degrees of reliability since they capture different sets of raw data; and
- sensors vary in their ability to capture data and generate relevant information (e.g., active sensors may be more reliable than passive sensors in detecting whether there is a device at a certain address).

If an unreliability rate for a sensor can be determined before it is used, it can be incorporated into the bpa reported by the sensor before the combination process commences. This can be done via the use of a *discounting method* or *importance weight*.

Based on a given unreliability rate, a discounting method in general decreases the mass of the focal elements, increasing ignorance. Various discounting methods exist, the most popular one proposed by Shafer [46] is given as:

$$\begin{aligned} m^*(A) &= tm(A), \forall A \subset \Omega \\ m^*(\Omega) &= 1 - t + tm(\Omega) \end{aligned} \tag{17}$$

where t $(0 \leq t \leq 1)$ is the unreliability rate of the source.

While a discounting method is quite efficient in reducing the impact of a piece of evidence produced by a sensor across all propositions, the method is not suitable where the reliability of a sensor varies according to *specific* propositions defined in the frame of discernment. In such a situation, a more appropriate method is to adjust the impact of a piece of evidence on individual propositions. Using a so-called *importance weight*, one can specify the reliability of a sensor for each proposition in the frame of discernment (e.g., sensor A is twice as reliable as sensor B in detecting proposition X, but is half as reliable as sensor B in detecting proposition Y). Fan and Zuo provide a full description of the use of an importance weight in [16].

If an unreliability rate cannot be determined beforehand (or if it is too hard to do so), it can be estimated in real-time based on the degree of conflict among pieces of evidence provided by the sensors. Dynamic detection of the reliability of sensors has been studied by several research groups. These research groups share the view that a high degree of conflict among pieces of evidence indicates unreliability of the sources of evidence, and that the sources that provide conflicting evidence should be considered unreliable and therefore should be discarded. This view is not valid in the context of Influx$_1$ where conflicting information is actually of interest and should be investigated rather than discarded. For this reason, while various discounting methods have been implemented in Influx$_1$, dynamic estimation of the reliability of sensors is not currently considered for implementation.

3.4 Independent evidence assumption and conflict distribution

The D-S rule requires pieces of evidence (in the form of belief functions) being combined to be distinct/independent. In addition, when pieces of evidence to be combined are conflicting, the D-S rule manages the situation by distributing the conflicting mass (i.e., the mass of the empty set) proportionally to the focal sets. Such a conflict management scheme may not be ideal for all contexts. For this reason, studies have been carried out which attempt to lift the independence assumption imposed on the pieces of evidence as well as to offer various ways to manage conflict, resulting in the numerous DS-related combination rules that can be found in the literature.

Tables 5 and 6 summarise a collection of combination rules studied in the literature, with the following desirable algebraic characteristics:

- **Commutativity**: $\forall m_1, \forall m_2 : m_1 \otimes m_2 = m_2 \otimes m_1$

  Commutativity implies that the order in which the evidence is combined does not affect the final result.

- **Associativity**: $\forall m_1, \forall m_2, \forall m_3$: $(m_1 \otimes m_2) \otimes m_3 = m_1 \otimes (m_2 \otimes m_3)$

  Associativity allows for (order-independent) pairwise computation, enabling local computation.

- **Quasi-associativity**: Associativity that is achieved by preserving certain intermediate results throughout the combination process.

Rule	Formula	Comm.	Asso.	Quasi-asso.	Idemp.
Conjunctive [51]	$m_{12}^{\cap}(A) = \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cap C = A}} m_1(B)m_2(C)$	yes	yes		
D-S [46]	$m_{12}^{DS}(A) = \frac{1}{1-K_{12}} \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cap C = A}} m_1(B)m_2(C)$	yes	yes		
Disjunctive [10]	$m_{12}^{\cup}(A) = \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cup C = A}} m_1(B)m_2(C)$	yes	yes		
Yager's [65]	$m_{12}^{Y}\Theta = m_1(\Theta)m_2(\Theta) + \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cap C = \emptyset}} m_1(B)m_2(C)$ $m_{12}^{Y}(A) = \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cap C = A}} m_1(B)m_2(C)$	yes		yes	
Averaging [35]	$m_{12}^{Avg}(A) = \frac{1}{2}\left[m_1(A) + m_2(A)\right]$	yes			yes
Dubois-Prade [11]	$m_{12}^{DP}(A) = \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cap C = A}} m_1(B)m_2(C) + \sum_{\substack{B,C \in 2^{\Theta}, \\ B \cup C = A, B \cap C = \emptyset}} m_1(B)m_2(C)$	yes		yes	

Table 5: Combination rules and their algebraic characteristics.

Rule	Formula	Comm.	Asso.	Quasi-asso.	Idemp.
PCR5 [48]	$$m_{12}^{PCR5}(A) = m_{12}^{\cap}(A) + \sum_{\substack{B\in 2^{\Theta},\\ B\cap A=\emptyset}} \left[\frac{m_1(A)^2 m_2(B)}{m_1(A)+m_2(B)} + \frac{m_2(A)^2 m_1(B)}{m_2(A)+m_1(B)}\right]$$	yes		yes	
Cumulative [28]	For $m_1(\Theta) \neq 0 \vee m_2(\Theta) \neq 0$: $$\begin{cases} m_C(A) = \frac{m_1(A)m_2(\Theta)+m_2(A)m_1(\Theta)}{m_1(\Theta)+m_2(\Theta)-m_1(\Theta)m_2(\Theta)}; \text{and} \\ m_C(\Theta) = \frac{m_1(\Theta)m_2(\Theta)}{m_1(\Theta)+m_2(\Theta)-m_1(\Theta)m_2(\Theta)} \end{cases}$$ for $m_1(\Theta) = 0 \wedge m_2(\Theta) = 0$: $$\begin{cases} m_C(A) = \gamma^1 m_1(A) + \gamma^2 m_2(A), \\ m_C(\Theta) = 0 \end{cases}$$ where $$\begin{cases} \gamma^1 = \lim_{m_1(\Theta)\to 0, m_2(\Theta)\to 0} \frac{m_2(\Theta)}{m_1(\Theta)+m_2(\Theta)}. \\ \gamma^2 = \lim_{m_1(\Theta)\to 0, m_2(\Theta)\to 0} \frac{m_1(\Theta)}{m_1(\Theta)+m_2(\Theta)} \end{cases}$$	yes	yes		
Cautious [7]	$$m_1 m_2 = \bigcap_{A\subset\Theta} A^{w_1(A)\wedge w_2(A)}$$ where $$w_{12}(A) = w_1(A) \wedge w_2(A), A \in 2^{\Theta}\backslash\Theta$$ and $$w(A) = \begin{cases} \frac{\Pi_{B\supseteq A,\lvert B\rvert\notin 2N} q(B)}{\Pi_{B\supseteq A,\lvert B\rvert\in 2N} q(B)} \text{if} \lvert A\rvert \in 2N, \text{and} \\ \frac{\Pi_{B\supseteq A,\lvert B\rvert\in 2N} q(B)}{\Pi_{B\supseteq A,\lvert B\rvert\notin 2N} q(B)} \text{otherwise.} \end{cases}$$	yes	yes		yes

Table 6: Combination rules and their algebraic characteristics (continue from Table 5).

- **Idempotence**: $\forall m_1$: $m_1 \otimes m_1 = m_1$

  Idempotence is important where there is non-distinct evidence since it prevents a combination rule from counting a piece of elementary evidence more than once.

The collection of combination rules can be analysed from a number of perspectives:

- With regards to the reliability of sources, most of the rules are based on two fundamental rules: the *conjunctive* rule (including the open-world conjunctive rule and closed-world D-S rule) and the *disjunctive* rule. A conjunctive rule looks for consensus between given bpas where the sources are sufficiently reliable, and a disjunctive rule tends to average out all the given information and should be used when not all the sources are reliable (e.g., one, or probably more, sources may be unreliable but one does not know which one).

  Here, one combination rule that is worth noticing is the cumulative rule which does not strictly belong to either the conjunctive or disjunctive category. The rule forms its own category; that is, the rule is applicable in situations where the pieces of evidence are *distinct* and *independent* but come from *identical* sources. In this case, the evidence is said to be combined in a cumulative fashion.

- In terms of conflict management, the different combination rules can be distinguished in particular according to the way they distribute conflict mass. Two main categories can be discerned: redistribution of global conflict mass (such as the D-S and Yager's combination rules) and redistribution of the partial conflict mass (such as the Dubois-Prade and PCR5 combination rules). For instance, the D-S and Yager's rules distribute the global conflict mass (the conflict mass between all the focal sets, if any, of the two combined bpas) to the focal sets and to the empty set, respectively. Unlike the D-S and Yager's rules, the Dubous-Prade and PCR5 rules distribute the partial conflict mass (between a group of focal sets) to the relevant subsets, and thus are supposed to be more precise when consideration of conflict is a focus.

- With respect to the types of evidence to be combined, the combination rules can also be discriminated based on the independence/dependence restriction imposed on the evidence. More specifically, the conjunctive, D-S, Yager's, Dubois-Prade, PCR5 and disjunctive rules assume the pieces of evidence to be combined are distinct and independent. Whereas for the averaging and cautious rules, pieces of evidence are required to be non-independent. To be able to deal with bodies of evidence that are not distinct and independent, the latter group of rules share one common algebraic property: they are all idempotent, preventing themselves from counting a piece of elementary evidence more than once.

Generally, the majority of combination rules discussed in the literature are sensitive to the independence/dependence assumption of evidence: applying an unsuitable combination rule to fuse evidence can yield inaccurate results. We have conducted an empirical comparison of the collection of combination rules presented in Tables 5 and 6. In general, the experimental results demonstrate the adequacy of the D-S rule as a combination (or fusion) operator to fuse independent pieces of evidence in various circumstances. More particularly, when receiving an appropriate treatment for dealing with highly conflicting evidence, the D-S rule exhibits behaviour similar to that of the proposed alternatives while maintaining mathematical simplicity and possessing highly-desirable characteristics such as commutativity and associativity.

4 The reasoning framework for Influx$_1$

D-S theory provides a mathematically sound approach for reasoning in situations where the information may be uncertain and incomplete. The extent to which the theory is realised in a practical setting naturally depends on the goals and characteristics of the specific reasoning problem to be addressed. In this regard, Influx is envisaged to be a reasoner handling real-life situations where the knowledge can be imperfect, the knowledge base potentially large and distributed, and efficiency (among other qualities) is of significant importance. As such, complete implementation of D-S theory is not a prime goal in the design and development of Influx; rather, the theory is utilised as a means to facilitate the development of appropriate mechanisms to represent and reason with uncertainty.

Nevertheless, realising D-S theory in practical applications is not straight-forward. Despite its rich literature, the amount of research work devoted to the application of D-S theory is fairly modest in comparison to the numerous efforts in the field dedicated to theoretical work. Two main reasons for this may be: (i) the heavy use of mathematical notation in the majority of work on D-S theory has hindered practitioners in the field of artificial intelligence, and (ii) the potentially high computational complexity associated with a straight-forward implementation of the D-S rule in the general case (being exponential to the size of the frame of discernment). Due to the latter reason, using the theory for uncertain reasoning in a practical context often necessitates the determination of an appropriate balance between the merits of the theory and the desired properties of the reasoning system to be implemented. To this end, Influx$_1$ aspires to achieve, as much as possible, the desired practical objectives while capitalising on the ideas and techniques offered by the D-S theory where suitable.

We have developed the first version of Influx, Influx$_1$, a simple nonmonotonic reasoning tool and framework which takes advantage of the D-S formalism to represent and reason with uncertainty. Conceptually, development of a reasoning framework is a matter of

- deciding on the representation of the relevant concepts (*concept representation*),
- determining the representation of the beliefs associated with the concepts (*belief representation*),
- specifying the mechanisms for belief combination, should there be more than one belief pertaining to a concept (*belief combination*), and
- defining the relationships between concepts which allows inferencing across the concepts to be carried out (*belief propagation*).

In this section we briefly review the reasoning framework of Influx$_1$ from such perspectives. Please note that design decisions and techniques pertaining to the implementation of Influx$_1$ are not discussed in this document.

For the purpose of illustration, Figure 5 graphically depicts a simple example of a rule network in Influx$_1$, various portions of which will be referred to in the discussions that follow. Before continuing further, it is important to clarify some of the concepts and terminology as follows:

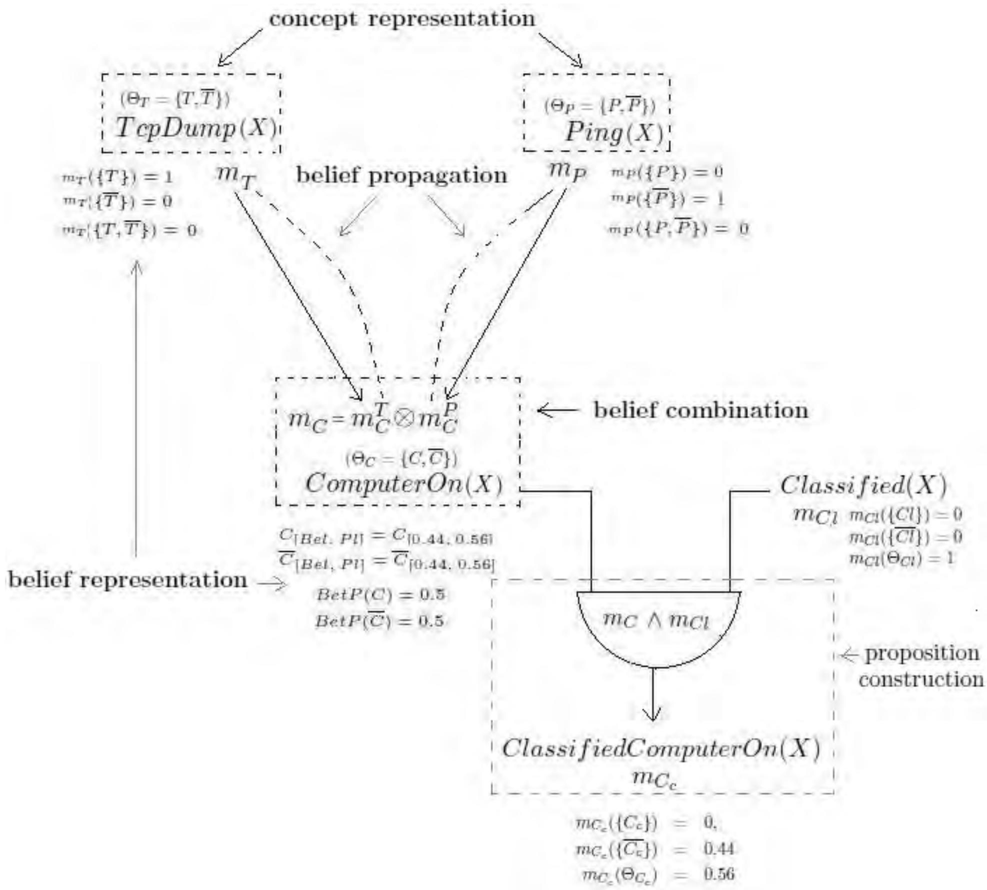


Figure 5: A simple rule network in Influx$_1$, illustrating concept representation, belief representation, belief combination and belief propagation. In this figure, $\otimes$ and $\wedge$ denote the 'fusion' and 'and' operators, respectively. The notation $\wedge$ will be defined in Section 4.1.1.

- $TcpDump(X)$, $Ping(X)$, $ComputerOn(X)$ and $ClassifiedComputerOn(X)$ are referred to as propositions. $TcpDump(X)$ denotes the proposition that 'the TcpDump sensor observes packets from computer X'. Similarly, $Ping(X)$ denotes the proposition that 'the Ping sensor receives responses from X', $ComputerOn(X)$ denotes the proposition that 'X is on', $Classified(X)$ denotes the proposition that 'X is classified', and $ClassifiedComputerOn(X)$ denotes the proposition that 'X is on and classified' whose truth value reflects the belief pertaining to $ComputerOn(X) \wedge Classified(X)$. Each proposition is associated with a binary frame of discernment (BFoD), e.g., $TcpDump(X)$ is associated with $\Theta_T = \{T, \overline{T}\}$ and $Ping(X)$ associated with $\Theta_P = \{P, \overline{P}\}$.

- The degree of certainty regarding the truth of each proposition is defined in the form of a bpa m, e.g., m_T, m_P and m_C for the propositions $TcpDump(X)$, $Ping(X)$ and $ComputerOn(X)$, respectively. Such a bpa expresses the belief mass assigned to all possible states of a proposition (e.g., $\{T\}$, $\{\overline{T}\}$ and $\Theta_T = \{T, \overline{T}\}$ in the case of $TcpDump(X)$). For example, the bpa m_T shows that the TcpDump sensor *certainly* observes packets from X ($m_T(T) = 1$). Besides bpas, the belief for a proposition, when required, can be represented in other ways, such as the belief function Bel, the plausibility function Pl and the pignistic function $BetP$. The pignistic function $BetP$ provides a probabilistic interpretation of the belief associated with a piece of knowledge which will be discussed in Section 4.2. For the sake of simplicity, when the distinction does not matter, all such functions (m, Bel, Pl and $BetP$) are referred to as belief functions (or simply beliefs).

- Propositions not directly observed (e.g., $ComputerOn(X)$) may be inferred from related propositions whose belief may be known (e.g., $TcpDump(X)$ and $Ping(X)$). Whether the TcpDump sensor observes packets from X, and whether the Ping sensor receives responses from X, can reveal some information about whether computer X is on. In this respect, $TcpDump(X)$ and $Ping(X)$ are generally referred to as information sources for $ComputerOn(X)$, and m_T and m_P are the beliefs associated with the information sources. The belief associated with $TcpDump(X)$ (m_T) can be combined with knowledge about the relation between $TcpDump(X)$ and $ComputerOn(X)$ to produce m_C^T, which is the belief about whether the computer X is on, given m_T. Likewise, we can produce m_C^P, which is the belief about whether the computer X is on, given m_P. We refer to the propagated beliefs m_C^T and m_C^P as *opinions* about $ComputerOn(X)$ provided by the respective information sources.

- The proposition $ComputerOn(X)$ has two independent opinions (m_C^T and m_C^P) about its status. To this end, the belief for $ComputerOn$ (m_C), is determined through the combination of those opinions ($m_C = m_C^T \otimes m_C^P$, where $\otimes$ denotes a conjunctive fusion operator).

4.1 Concept representation

Concepts in the knowledge base of Influx$_1$ may be simple or complex. For instance, a concept can be an elementary piece of knowledge such as

```
ComputerActive,
```

the knowledge that a computer is active. A concept can also be structured with embedded contextual information such as

```
Computer.Software.AntiVirusInstalled(fileserver, SuperAV),
```

the knowledge that a particular computer ($fileserver$) has antivirus software installed ($SuperAV$) where $AntiVirusInstalled$ carries the semantics defined within the context $Computer.Software$. Within Influx$_1$, this structured representation (or structured ID) allows knowledge and relationships to be described with a flexible and informal form of ontology and enables certain techniques for efficient computation and memory use.

As the notion of frames of discernment (FoDs) is fundamental for knowledge representation in D-S theory, all concepts must be defined with respect to one or more FoDs. To this end, the building blocks for representing knowledge in Influx$_1$ are the so-called binary frames of discernment (BFoDs). Each BFoD is associated with a concept, and contains two elements: the concept and its negation (e.g., $\Theta_A = \{A, \bar{A}\}$).

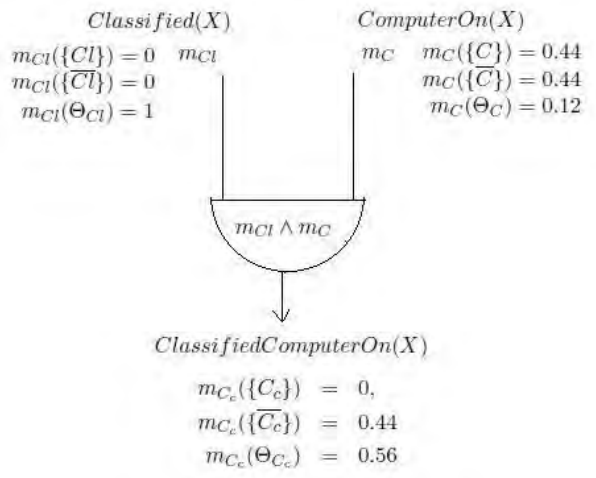


Figure 6: An illustration of a proposition constructed from other propositions using the 'and' connective.

We group concepts according to their inter-relationships, which is important for the devising and determining of the appropriate mechanisms to combine their beliefs. The first group refers to those concepts which are not mutually exclusive. This group can be further divided into two subgroups:

- *independent concepts* whose beliefs are not correlated (i.e., change in the belief of one concept has no (or negligible) effect on the belief of the others), and
- *dependent concepts* whose beliefs are correlated (i.e., change in the belief of one concept has effect on the belief of the others).

The second group refers to those concepts which are *mutually exclusive*. The BFoDs associated with the concepts in this case are considered specific *coarsenings* (see Section 3) of an implicit frame Ω which includes all such concepts as part or all of its elements.

With respect to belief representation (which will be further discussed in the next section), each concept A is associated with a bpa m_A specifying the degree of belief for $\{A\}$, $\{\overline{A}\}$ and Θ_A.[12] Manipulation of the concept's belief can be mathematically performed in a manner similar to that of a logical proposition with infinite truth values, *albeit* with different interpretations[13]. More specifically, a concept, hereby referred to as a proposition, can be transformed into, or inferred from, a related one; or can be built from more primitive propositions. For instance, $ClassifiedComputerOn(X)$ can be built from $Classified(X)$ and $ComputerOn(X)$ using the logical connective *and* as shown in Figure 6. The following discussion is concerned with combining beliefs associated with different propositions to construct a new proposition. Propagation of belief from one proposition to another will be treated in Section 4.4.

[12]Hereafter, for the sake of simplicity of notation, we at times do not distinguish between elements and singleton subsets of a set when it is clear by the context.

[13]We are dealing with infinite degrees of belief about the truth of a proposition rather than infinite truth values of the proposition.

4.1.1 Proposition construction

Proposition construction is essentially concerned with the combination of beliefs expressed on *different* frames of discernment. More specifically, a proposition (e.g., B) can be built from a set $\mathbf{A}$ of n other propositions (e.g., $A_1, A_2, \ldots, A_n$). This is achieved by computing the belief of B (i.e., m_B) as the combination of the beliefs expressed on Θ_{A_i} (i.e., $m_{A1}, m_{A2}, \ldots, m_{An}$). In Influx$_1$, such combination is performed according to the fundamental criteria of necessity or sufficiency:

- *Necessity* refers to situations where all elements in $\mathbf{A}$ must be confirmed in order for B to be confirmed. Broadly speaking, the belief value for B is only high when the belief values for all A_i are high.

- *Sufficiency*, in contrast, refers to situations where any element in $\mathbf{A}$ being confirmed will result in B being confirmed. Broadly speaking, the belief value for B is high when the belief value for any A_i is high.

The combination process will reflect the criteria relevant to the concepts, and the type of relationship they share (independent, dependent or mutually exclusive). To this end, a number of functions addressing proposition construction have been derived for implementation in Influx$_1$. The derived functions are depicted in Table 7 and elaborated below.

4.1.1.1 Combination of independent propositions Given a set of propositions A_i with an independent relationship, the belief for a proposition B can be computed by combining the beliefs associated with A_i using the logical connectives *and* (i.e., $B = A_1 \wedge A_2 \wedge \ldots \wedge A_n$ if A_i are necessary criteria for B) or *or* (i.e., $B = A_1 \vee A_2 \vee \ldots \vee A_n$ if A_i are sufficient criteria for B). To facilitate reasoning in Influx$_1$, such connectives, among others, must be redefined to take into account the adopted D-S belief representation. Multiple efforts to integrate logical reasoning and Dempster-Shafer belief functions have been made (such as [9], [17], [15] and [18] in the context of information fusion, belief maintenance systems and non-monotonic deductive reasoning). However the formulation of the logical connectives in the mentioned works (particularly [17]) assumes a global frame of discernment on which the bpas of the constituent propositions are defined, and thus cannot be applied directly in this context.

To this end, it is necessary to derive the formulas for the *and* and *or* connectives for any two propositions A_1 and A_2. This may be achieved by first computing the associated belief distribution on the Cartesian product $\Theta_{(A_1,A_2)}$ (where Θ_{A1} and Θ_{A2} are the frames of discernment pertaining to A_1 and A_2, respectively). Such a belief distribution (see Table 8) allows belief associated with a conjunction between every subset of Θ_{A1} with that of Θ_{A2} to be captured. Applying Kleene's truth tables for *and* and *or* for three-valued logic (see Figure 7) to the computed belief distribution then enables a derivation of the formulas for the logical connectives (or operators) as given below.

Relationship between A_1, A_2 and B	Operator type	Relationship between A_1 and A_2	Specific operator	Influx$_1$ syntax	Formula
Sufficiency	$B = or(A_1, A_2)$	independent	$\vee$	`OR`	$m(B) = m(A_1) + m(A_2) - m(A_1)m(A_2)$ $m(\overline{B}) = m(\overline{A_1})m(\overline{A_2})$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$
		dependent	$\vee^{sub}$	`MAX`	$m(B) = max(m(A_1), m(A_2))$ $m(\overline{B}) = min(m(\overline{A_1}), m(\overline{A_2}))$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$
		mutually exclusive	$\vee^{add}$	`ADD`	$m(B) = m(A_1) + m(A_2)$ $m(\overline{B}) = \frac{1}{2}(m(\overline{A_1}) + m(\overline{A_2}) - m(A_1) - m(A_2))$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$
Necessity	$B = and(A_1, A_2)$	independent	$\wedge$	`AND`	$m(B) = m(A_1)m(A_2)$ $m(\overline{B}) = m(\overline{A_1}) + m(\overline{A_2}) - m(\overline{A_1})m(\overline{A_2})$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$
		dependent	$\wedge^{sub}$	`MIN`	$m(B) = min(m(A_1), m(A_2))$ $m(\overline{B}) = max(m(\overline{A_1}), m(\overline{A_2}))$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$
Neither necessity nor sufficiency	$B = mean(A_1, A_2)$	independent/ dependent	$mean$	`MEAN`	$m(B) = \frac{1}{2}(m(A_1) + m(A_2))$ $m(\overline{B}) = \frac{1}{2}(m(\overline{A_1}) + m(\overline{A_2}))$ $m(\Theta_B) = 1 - m(B) - m(\overline{B})$

Table 7: Combination operators implemented in Influx$_1$ and their notations used in this document. In the table, the proposition B is constructed from the propositions A_1 and A_2 according to the the criteria of sufficiency and necessity.

The *and* operator Let B denote a proposition that reflects the *conjunctive truth* of the two propositions A_1 and A_2 ($B = A_1 \wedge A_2$), the belief function associated with B is computed as:

$$\begin{aligned} m(B) &= m(A_1)m(A_2), \\ m(\overline{B}) &= m(\overline{A_1})m(A_2) + m(\overline{A_1})m(\overline{A_2}) + m(\overline{A_1})m(\Theta_{A2}) + m(\overline{A_2})m(A_1) + m(\overline{A_2})m(\Theta_{A1}) \\ &= m(\overline{A_1}) + m(\overline{A_2}) - m(\overline{A_1})\, m(\overline{A_2}), \\ m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \end{aligned} \tag{18}$$

For instance, the proposition $ClassifiedComputerOn(X)$ in Figure 6 is considered confirmed only if both $Classified(X)$ and $ComputerOn(X)$ are confirmed, and it is not confirmed otherwise. Given m_{Cl} and m_C as depicted in the figure and using the above formula, one obtains the following belief computed for $ClassifiedComputerOn(X)$:

$$\begin{aligned} m_{C_c}(C_c) &= 0 \times 0.44 = 0, \\ m_{C_c}(\overline{C_c}) &= 0.44 + 0 - 0.44 \times 0 = 0.44, \\ m_{C_c}(\Theta_{C_c}) &= 1 - 0 - 0.44 = 0.56. \end{aligned}$$

	A_2	$\overline{A_2}$	Θ_{A_2}
A_1	$m(A_1)m(A_2)$	$m(A_1)m(\overline{A_2})$	$m(A_1)m(\Theta_{A_2})$
$\overline{A_1}$	$m(\overline{A_1})m(A_2)$	$m(\overline{A_1})m(\overline{A_2})$	$m(\overline{A_1})m(\Theta_{A_2})$
Θ_{A_1}	$m(\Theta_{A_1})m(A_2)$	$m(\Theta_{A_1})m(\overline{A_2})$	$m(\Theta_{A_1})m(\Theta_{A_2})$

Table 8: A belief distribution on $\Theta_{(A_1,A_2)}$ computed from the belief functions associated with Θ_{A_1} and Θ_{A_2}.

and	**T**	**F**	**U**
T	T	F	U
F	F	F	F
U	U	F	U

or	**T**	**F**	**U**
T	T	T	T
F	T	F	U
U	T	U	U

Figure 7: Three-valued logic truth tables for 'and' and 'or', where U represents 'unknown'.

The *or* operator Let B denote a proposition that reflects the *disjunctive truth* of the two propositions A_1 and A_2 ($B = A_1 \vee A_2$), the belief function associated with B is given by:

$$\begin{aligned} m(B) &= m(A_1)m(A_2) + m(A_1)m(\overline{A_2}) + m(A_1)m(\Theta_{A_2}) + m(A_2)m(\overline{A_1}) + m(A_2)m(\Theta_{A_1}) \\ &= m(A_1) + m(A_2) - m(A_1)m(A_2), \\ m(\overline{B}) &= m(\overline{A_1})\, m(\overline{A_2}), \\ m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \end{aligned} \tag{19}$$

For example, let $Vulnerable(X)$ correspond to $Unpatched(X) \vee DisabledFirewall(X)$, which means the proposition $Vulnerable(X)$ is considered confirmed if $Unpatched(X)$ or $DisabledFirewall(X)$ is confirmed. Using the above formula, the bpa associated with V

can be computed as

$$\begin{aligned} m_V(V) &= 0.44 + 0 - 0.44 \times 0 = 0.44, \\ m_V(\overline{V}) &= 0.44 \times 0 = 0, \\ m_V(\Theta_V) &= 1 - 0.44 - 0 = 0.56, \end{aligned}$$

where V, U and D denote $Vulnerable(X)$, $Unpatched(X)$ and $DisabledFirewall(X)$, respectively, and

$$m_U(U) = 0, m_U(\overline{U}) = 0, m_U(\Theta_U) = 1,$$

$$m_D(D) = 0.44, m_D(\overline{D}) = 0.44, m_D(\Theta_D) = 0.12.$$

As presented above, the *and* and *or* operators are used when the set **A** provides necessary or sufficient criteria for B, respectively. In a number of situations, $A_1, A_2, \ldots, A_n$ are criteria for B but each is neither sufficient nor necessary. That is, the confirmation of any single A_i is not sufficient to lead to the confirmation of B, but at the same time, it is not necessary for all A_i to be confirmed in order for B to be confirmed. To compute the belief for B in this case, the *mean* operator is used which does not rely on the confirmation or otherwise of any one criteria, but takes into account all of the criteria and 'balances out' their respective beliefs.

The *mean* operator Let B denote a proposition that reflects the overall truth of the two propositions A_1 and A_2, the belief function associated with C is given by:

$$\begin{aligned} m(B) &= \frac{1}{2}(m(A_1) + m(A_2)) \\ m(\overline{B}) &= \frac{1}{2}(m(\overline{A_1}) + m(\overline{A_2})), \\ m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \end{aligned} \tag{20}$$

For example, $FastCPU(X)$ and $BigHardDrive(X)$ are criteria for the proposition $UsefulComputer(X)$. Each of the criteria on its own may not be sufficient for the computer X to be useful, nor is the absence of one necessary to render the computer not useful. Using the *mean* operator, the bpa associated with $UsefulComputer(X)$ can be computed as

$$\begin{aligned} m_U(U) &= \frac{1}{2}(0.44 + 0) = 0.22, \\ m_U(\overline{U}) &= \frac{1}{2}(0.44 + 0) = 0.22, \\ m_U(\Theta_U) &= 1 - 0.22 - 0.22 = 0.56. \end{aligned}$$

where U, F and B denote $UsefulComputer(X)$, $FastCPU(X)$ and $BigHardDrive(X)$, respectively, and

$$m_F(F) = 0, m_F(\overline{F}) = 0, m_F(\Theta_F) = 1,$$

$$m_B(B) = 0.44, m_B(\overline{B}) = 0.44, m_B(\Theta_B) = 0.12.$$

When the criteria A_i have varying degree of importance to the constructed proposition, this can be captured by incorporating a relative importance weight to each A_i (see Section 4.3.1.1) before performing the *mean* operation.

4.1.1.2 Combination of dependent propositions Suppose that there exist explicit/implicit relations among the propositions A_i, and thus their beliefs are correlated. For instance, the two propositions $InternetEnabled(X)$ and $EmailActive(X)$ can be considered dependent since the belief of one proposition can be inferred from the other. In this case, the *mean* operator remains applicable due to its idempotency (i.e., the belief of a proposition is unchanged when combined with itself). However, using the previously derived *and* and *or* functions (which treat the belief associated with each A_i as an independent contribution to the combined belief) would incorrectly amplify the belief computed for B. To address combination in such a situation, Influx$_1$ offers cautious forms of the *and* and *or* functions which assume a (positive) correlation of belief among the propositions being combined.

The *subsumptive and* operator Let the set $\mathbf{A}$ represent necessary criteria A_i for B, where the beliefs associated with A_i are correlated (and thus may be considered to provide overlapping or duplicate information). In this case, it is plausible for the belief of B to be computed according to the criteria that are least confirmed (those with the minimum belief and/or the minimum plausibility). Translated to D-S belief representation, the belief function associated with B is computed as:

$$\begin{aligned}
m(B) &= min(m(A_1), m(A_2)), \\
m(\overline{B}) &= 1 - Pl(B) \\
&= 1 - min(Pl(A_1), Pl(A_2)) \\
&= 1 - min(1 - m(A_1), 1 - m(A_2)), \\
&= max(m(A_1), m(A_2)), \\
m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \qquad (21)
\end{aligned}$$

The derived function mathematically resembles the *weak conjunction* of Łukasiewicz's many-valued logic [33]. We refer to this function as a *subsumptive and* (or and^{sub}, or $\wedge^{sub}$).

The *subsumptive or* operator Let the set $\mathbf{A}$ represent the sufficient criteria A_i for B, where the beliefs associated with A_i are correlated (and thus may be considered to provide overlapping or duplicate information). In this case, it is plausible for the belief of B to be computed according to the criteria that are most confirmed (those with the maximum belief and/or the maximum plausibility). Translated to D-S belief representation, the belief function associated with B is computed as:

$$\begin{aligned}
m(B) &= max(m(A_1), m(A_2)), \\
m(\overline{B}) &= 1 - Pl(B) \\
&= 1 - max(Pl(A_1), Pl(A_2)) \\
&= 1 - max(1 - m(A_1), 1 - m(A_2)), \\
&= min(m(A_1), m(A_2)), \\
m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \qquad (22)
\end{aligned}$$

The derived function mathematically corresponds to the *weak disjunction* of Łukasiewicz's many-valued logic [33]. We refer to this function as a *subsumptive or* (or or^{sub}, or $\vee^{sub}$).

4.1.1.3 Combination of mutually exclusive propositions Suppose the propositions A_i are mutually exclusive (e.g., $Restricted(X)$, $Secret(X)$ and $TopSecret(X)$), A_i are considered to correspond to elements in a common implicit frame Ω, and the BFoD associated with each A_i is a coarsening of Ω focusing on A_i. As such, the beliefs associated with A_i can now be expressed on the same space of discernment 2^{Ω}, thereby enabling direct manipulation of the beliefs by means of set-theoretical operations on the discernment space. Assuming the mutually exclusive relationship between A_i is appropriately captured (thus, it is always that $\sum_i(m(A_i)) \leq 1$ and $m(\overline{A_i}) \geq \sum_{j\neq i} m(A_j)$), proposition combination can be performed as given below.

The *additive or* operator Let the set $\mathbf{A}$ represent sufficient criteria A_i for B, where A_i are mutually exclusive. In this particular case, the proposition $B = A_1 \vee A_2 \vee \ldots \vee A_n$ corresponds to the disjunctive set $\{A_1, A_2, \ldots, A_n\}$ defined on Ω, and its belief can be approximately computed as follows:

$$\begin{aligned} m(B) &= m(A_1) + m(A_2), \\ m(\overline{B}) &= \frac{m(\overline{A_1}) + m(\overline{A_2}) - m(A_1) - m(A_2)}{2}, \\ m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \end{aligned} \tag{23}$$

The above formula performs exact computation when $m(\overline{A_i}) = \sum_{j\neq i} m(A_j)$.

A conjunction of mutually exclusive propositions would lead to a contradiction in classical logics, or an empty set in set theory. As such, constructing B when A_i serve as necessary criteria, or neither necessary or sufficient criteria, for B is often not of practical use. Therefore, the implementation of Influx$_1$ does not include functions to perform such operations.

The formulas for the different versions of *and* and *or* presented above are both associative and commutative, and thus can be efficiently applied in a pairwise fashion to an arbitrary number of propositions. When the beliefs associated with A and B are certain and A and B are independent, the formulas simplify to the classical *and* and *or* connectives. Unlike the *and*- and *or*-typed operators, the *mean* operator is only quasi-associative — in Influx$_1$ this is resolved by ensuring the operation is performed in an associative manner, and thus can be utilised in the same way as other operators.

4.1.2 General discussion

As emphasised at the beginning of the document, Influx$_1$ does not aim to faithfully implement D-S theory, but to capitalise on the ideas and techniques offered by the theory to represent belief and reason under uncertainty. To this end, the rationale for utilising BFoDs in Influx$_1$ is multifold, including:

- this restricted form of FoD makes possible the derivation of simplified versions of various combination operators and propagation functions (see Sections 4.3 and 4.4), thus dramatically reducing the computational complexity of the reasoning process;

- this concept representation also directly facilitates logical reasoning, and significantly promotes other practically desirable properties (such as uniformity, scalability, distributivity and supporting changes in the knowledge structure).

Despite the mentioned advantages, one implication of the use of BFoDs is that the *full* expressiveness and reasoning power offered by D-S theory is partially compromised. Nevertheless, the theoretically appealing expressiveness offered by the D-S formalism is also the very reason for the impracticability of many applications based on D-S theory in the real world — recall that the computational complexity of the D-S rule is exponential to the size of the frame of discernment. It is thus a common practice to impose certain restrictions on the expressiveness of the formalism so as to achieve the practical objectives of a reasoning system. To this end, it is necessary to speculate on the degree to which the expressiveness of the D-S caculus can be preserved with the use of BFoDs as primitive building blocks for knowledge representation in Influx_1.

In Influx_1, a frame Ω with an arbitrary size can be represented by means of one or more BFoDs. Such BFoDs correspond to various coarsenings of Ω each focusing on a proposition of interest. In such cases, while incompleteness pertaining to the belief associated with each proposition (i.e., the ignorance value) remains intact, a *direct* expression of incomplete knowledge over a set of mutually exclusive propositions is no longer possible (as a bpa cannot assign a single belief mass to a set of propositions across different frames). Also, implicit mutual-exclusion or set-inclusion relationships between such propositions is no longer implicitly handled by the D-S rule. If and as required, what is no longer directly represented or implicitly captured (with the use of a complex frame of discernment) may instead be represented by multiple propositions (each associated with a BFoD as a coarsening of Ω) with explicit relationships. This effectively allows one to model the useful portions of the discernment space (2^{Ω}) relevant to the problem at hand.

For the purpose of illustration, given the three propositions $Ubuntu(X)$, $RedHat(X)$, and $Linux(X)$, the BFoDs associated with the propositions can be considered coarsenings of an implicit frame $OS = \{Ubuntu, RedHat, Windows\_XP, Windows\_NT\}$ (assuming that OS is exhaustive). Here, $Ubuntu$, $RedHat$ and $Linux$ correspond to $\{Ubuntu\}$, $\{RedHat\}$ and $\{Ubuntu, RedHat\}$, respectively, and $\overline{Ubuntu}$, $\overline{RedHat}$ and $\overline{Linux}$ correspond to $\{RedHat, Windows\_XP, Windows\_NT\}$, $\{Ubuntu, Windows\_XP, Windows\_NT\}$ and $\{Windows\_XP, Windows\_NT\}$, respectively. The frame OS effectively allows the relationship between the propositions (i.e., the mutual exclusion and set-inclusion) to be discerned and implicitly captured within the frame. However without expressing the propositions on the common frame OS, such relationships need to be explicitly specified. For instance, the set-inclusion relationship between $Ubuntu(X)$ and $Linux(X)$:

$$Ubuntu(X) \overset{is\_a}{\rightarrow} Linux(X)$$

(i.e., $Ubuntu(X) \overset{1}{\rightarrow} Linux(X)$, $\overline{Ubuntu(X)} \overset{0}{\rightarrow} \overline{Linux(X)}$), may be expressed within Influx_1 using the *include* operator (`INC`) as

```
Linux(X) = INC(Ubuntu(X)).
```
[14]

The function specified above can be interpreted as 'if X is *Ubuntu*, it is *certainly Linux*; otherwise it is unknown if X is *Linux*'. Likewise, given the belief of $Ubuntu(X)$ is known, the belief for $RedHat(X)$ can be explicitly inferred as

$$Ubuntu(X) \overset{is\_not\_a}{\rightarrow} RedHat(X)$$

(i.e., $Ubuntu(X) \overset{1}{\rightarrow} \overline{RedHat(X)}$, $\overline{Ubuntu(X)} \overset{0}{\rightarrow} RedHat(X)$), and computed using the following Influx$_1$ syntax:

$$\texttt{RedHat(X)} = \texttt{NOT\^{}INC(Ubuntu(X))}. \tag{24}$$

In the same vein, if $RedHat(X)$ is known, the belief for $Ubuntu(X)$ can be explicitly represented and computed in Influx$_1$:

$$\texttt{Ubuntu(X)} = \texttt{NOT\^{}INC(RedHat(X))}. \tag{25}$$

The two functions above indicate that if X is *RedHat*, it is certainly not *Ubuntu* (otherwise, it is unknown if X is *Ubuntu*), and if X is *Ubuntu*, it is certainly not *RedHat* (otherwise, it is unknown if X is *RedHat*).

It is important to note that when $Ubuntu(X)$ and $RedHat(X)$ both receive evidence and transfer belief to each other (thus, necessitating that both functions (24) and (25) be invoked), one faces the problem of circular inferencing (as depicted in Figure 8a). Since Influx$_1$ does not yet handle bidirectional reasoning, circular inferencing can be problematic. To this end, circular inferencing can be avoided at the network level by means of 'divorcing' elements of a set of mutually exclusive propositions $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$. This can be done via the use of a dummy proposition A'_i for each A_i which collects and combines all the beliefs induced on A_i (excluding those induced from A_j, $j \neq i$) before transferring the combined belief to A_i and, at the same time, acting as evidence to disconfirm A_j ($A_j \in \mathbf{A}, j \neq i$). As shown in Figure 8b, provided that the sources (e and e') that give evidence for $isRedHat(X)$ and $isUbuntu(X)$ are independent, the multiple beliefs associated with $RedHat(X)$ and $Ubuntu(X)$ can be computed using a fusion operator (such as the D-S rule):

```
RedHat(X) = DS(isRedHat(X), NOT^INC(isUbuntu(X)))
Ubuntu(X) = DS(isUbuntu(X), NOT^INC(isRedHat(X))),
```

If the beliefs propagated to $isRedHat(X)$ and $isUbuntu(X)$ originate from overlapping sources, the method that combines the beliefs should appropriately reflect such dependency. Details relevant to the above combination and propagation functions are discussed in Sections 4.3 and 4.4 of the document.

As shown above, the discernment space 2^{Ω} can be potentially represented by means of BFoDs and explicit relationships between them. Yet, the presented approach has the following major implications:

[14] If the rule is intended to be applied to *any* computer X, the syntax should read: `Linux($X) = INC(Ubuntu($X))`.

Figure 8: The circular inferencing between $Ubuntu(X)$ and $RedHat(X)$ in (a) can be resolved by decoupling the two propositions via the use of the dummy propositions $isUbuntu(X)$ and $isRedHat(X)$ as shown in (b).

- An explicit representation of a portion of the discernment space by means of coarsened frames and the relationships between them (instead of having them implicitly embedded in Ω in the standard approach) generally results in a less compact network, as a trade-off for attaining the aforementioned practical goals of the reasoner. If Ω contains many elements, and the portion of interest, S, of the discernment space is large and complex, an explicit representation of the space of interest can be cumbersome. Fortunately, in many practical cases, S is relatively small and hierarchically structured as illustrated in Figure 4 and discussed in Section 3.2.3. Indeed in such situations, the use of a small number of binary frames (as coarsened frames of Ω), instead of a large frame of discernment, can provide a more intuitive and transparent way to capture the expert's knowledge, as well as eliminate the need to specify the complex relationship between any two frames with an arbitrary size.
- The transfer of belief between subsets in Ω as carried out by the D-S rule does not assume any direction. As Influx_1 does not yet handle bidirectional reasoning, inferencing between the propositions in S is carried out according to the direction explicitly specified, in a uniform manner with the rest of the rule network.

Both of the issues presented above are part of the agenda to be addressed in the design and implementation of Influx_2.

4.2 Belief representation

In Section 3, we presented three basic belief functions to capture imperfect knowledge expressed on a frame of discernment: basic probability assignment (bpa or m), belief function (Bel) and plausibility function (Pl). These three functions have direct correspondence and can be converted between each other through linear transformation (i.e, Möbiũs transform). For this reason, when the distinction between them does not matter, we will, hereafter, simply refer to them as belief functions.

Due to their mathematical equivalence, any of these functions can, in principle, be used to encode the knowledge base, and which function to use depends on preference and convenience. For instance, the function m can be recovered from Bel as

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \, Bel(B) \tag{26}$$

where $|\,.\,|$ denotes the cardinality.

Since a BFoD is used for concept representation in Influx_1, it is straight-forward to use the bpa m for storing and computing knowledge. However, Influx_1 supports these three belief functions, and the user may choose on each occasion which to use. Since a bpa m expresses belief on a frame of discernment at a primitive level, and thus enables direct manipulation of the belief mass assigned to individual propositions, it is a natural choice for describing input knowledge, and the current state of knowledge. Whereas for knowledge outcomes, the natural choice is the degree of belief and plausibility of propositions as supported by evidence.

As BFoDs (e.g., $\Theta_A = \{A, \overline{A}\}$) are the building blocks for representing knowledge in Influx$_1$, a bpa m associated with each concept is, in this particular case, defined by the three numbers $m(\{A\})$, $m(\{\overline{A}\})$, and $m(\Theta)$ (or $m(\{A, \overline{A}\})$). Then the degree of belief of A is simply given by:

$$Bel(\{A\}) = m(\{A\})$$

while the plausibility of A is computed as:

$$Pl(\{A\}) = m(\{A\}) + m(\Theta_A) = 1 - Bel(\{\overline{A}\})$$

and the quantity $m(\Theta_A)$ corresponds to the amount of ignorance.

The use of BFoDs and their associated bpas in Influx$_1$ provides the reasoning system a simple and uniform belief representation (i.e., the belief associated with every concept is represented in the form $(m(\{A\}), m(\{\overline{A}\})$, and $m(\Theta_A))$, contributing significantly to the efficiency of system development and deployment.

4.2.1 Belief and plausibility functions

In general, a belief function is useful in indicating the degree to which a proposition is believed to be true (due to the available evidence confirming/disconfirming it, and the defined relationships between the propositions), while the plausibility function suggests the extent to which a proposition could possibly be true (due to the same evidence and relationships). Commonly accepted interpretations for the associated interval $[Bel, Pl]$ are as follows:

- $A_{[0,\,1]}$: no knowledge at all about A.
- $A_{[0,\,0]}$: A is false.
- $A_{[1,\,1]}$: A is true.
- $A_{[0.25,\,1]}$: evidence provides partial support for A.
- $A_{[0,\,0.85]}$: evidence provides partial support for $\overline{A}$.
- $A_{[0.25,\,0.85]}$: the evidence simultaneously provides support for both A and $\overline{A}$.

4.2.2 Pignistic and ternary logic functions

Besides the belief and plausibility functions which are the two standard criteria for decision making, Influx$_1$ also provides two additional belief representations using the pignistic function $BetP$ and the ternary logic function T. The pignistic function endows users with the flexibility to adopt a probabilistic interpretation of knowledge (which is necessary when decision making is to be carried out based on *how probable a hypothesis is*), as well as the capability to switch to probabilistic reasoning at any point throughout the reasoning process. Likewise, the ternary logic function allows reasoning in Influx$_1$ to simplify into reasoning with three-valued logic (*true*, *false*, or *unknown*) when necessary.

Pignistic function In the framework of transferable belief models (TBMs), Smets [51] distinguished two different ways to quantify beliefs according to the mental level in which they are used: at the *credal level* (where beliefs are entertained), it is essential that beliefs are represented as belief functions; and at the *pignistic level* (where beliefs are used to make decisions), it is necessary that beliefs are represented as probability functions. To this end, the $BetP$ function allows one to transform a belief measure into a probability measure as defined below:

$$\begin{aligned} BetP(\theta_i) &= \sum_{\theta_i \in \theta \subseteq \Theta} \frac{m(\theta)}{|\theta|}, \\ BetP(\theta) &= \sum_{\theta_i \in \theta} BetP(\theta_i), \theta \subseteq \Theta. \end{aligned} \tag{27}$$

In the case of Influx$_1$ with BFoDs, this simplifies to the equal distribution of $m(\Theta_A)$ to $m(\{A\})$ and $m(\{\overline{A}\})$.

Example 7 Given the bpa m_{C_c} which expresses the belief associated with the proposition $ClassifiedComputerOn(X)$ (denoted as C_c) in Figure 9:

$$\begin{aligned} m_{C_c}(\{C_c\}) &= 0, \\ m_{C_c}(\{\overline{C_c}\}) &= 0.44, \\ m_{C_c}(\Theta_{C_c}) &= 0.56, \end{aligned}$$

this belief can be represented with Bel, Pl and $BetP$ as:

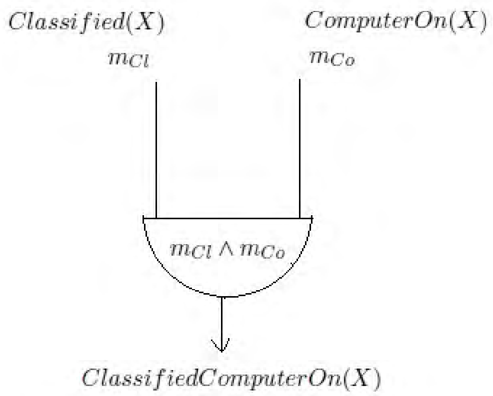
$$C_{c\,[Bel,\,Pl]} = C_{c\,[0,\,0.56]},$$

and

$$\begin{aligned} BetP(C_c) &= 0.28, \\ BetP(\overline{C_c}) &= 0.72. \end{aligned}$$

According to the $BetP$ function, the probability for the fact that 'a computer X is on and classified' is 28%. However, the Bel and Pl functions provides more insight into this result: (i) this fact is currently not confirmed by the available evidence ($m_{C_c}(\{C_c\}) = Bel_{C_c}(\{C_c\}) = 0$) while there is evidence that partially disconfirms this fact ($m_{C_c}(\{\overline{C_c}\}) = Bel_{C_c}(\{\overline{C_c}\}) = 0.44$); and (ii) in general this result is not strongly supported (indicated by a reasonably large amount of belief mass assigned to ignorance: $m_{C_c}(\Theta_{Cc}) = 0.56$) due to either a lack of evidence or a weak relation defined between the proposition and those from which it infers belief.

By supporting multiple modalities for belief representation, Influx$_1$ provides a rich and informative interpretation of a piece of knowledge. For instance, Figure 10 depicts two very different situations which give rise to the same answer in terms of probabilistic interpretation. In Figure 10a, the two sensors are not activated, and thus are completely



$$\begin{aligned} m_{C_c}(\{C_c\}) &= 0, \\ m_{C_c}(\{\overline{C_c}\}) &= 0.44, \\ m_{C_c}(\Theta_{C_c}) &= 0.56 \end{aligned} \qquad C_{c\,[Bel,\,Pl]} = C_{c\,[0,\,0.56]} \qquad \begin{aligned} BetP(C_c) &= 0.28 \\ BetP(\overline{C_c}) &= 0.72 \end{aligned}$$

Figure 9: An illustration of different ways to represent belief for $ClassifiedComputerOn(X)$ in Example 7. The bpa m_{C_c} (left) is generally suitable for capturing the input and current state of knowledge, while the other two representations (middle and right) are useful to summarise the outcome (specifically, as degrees of belief and plausibility using the Bel and Pl functions, or as probabilities using the $BetP$ function).

ignorant about whether there is any traffic or ping response from computer X (revealed in the vacuous bpas m_T and m_P), whereas in Figure 10b, $TcpDump$ certainly observes traffic from X but $Ping$ has not received any response from the computer (revealed in their total confirmation and disconfirmation bpas). In both cases, the computer X has an equal probability to be on or off (as indicated by the $BetP$ functions). However, the information provided by the Bel and Pl functions faithfully reflects the reality: (i) though it is totally plausible, there is absolutely no evidence informing the status of X in the former scenario (i.e., $C_{[0,\,1]}$ and $\overline{C}_{[0,\,1]}$), and (ii) there is an equal amount of evidence that supports both possible states of the computer X (i.e., $C_{[0.44,\,0.56]}$ and $\overline{C}_{[0.44,\,0.56]}$) in the latter scenario.

Ternary logic function Instead of simplifying reasoning with belief functions into probabilistic reasoning via the use of the pignistic function, the ternary logic function T effectively reduces reasoning with belief functions to three-valued logical reasoning in a seamless manner when the belief supporting a proposition reaches a sufficient degree of certainty defined by δ. The function T for a BFOD Θ is defined as:

$$\forall A \subset \Theta, T(A) = \begin{cases} 1 & \text{if } m(A) \geq \delta, \text{ where } \delta \in [0.5, 1] \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

$$T(\Theta) = 1 - \sum_{A \subset \Theta} T(A)$$

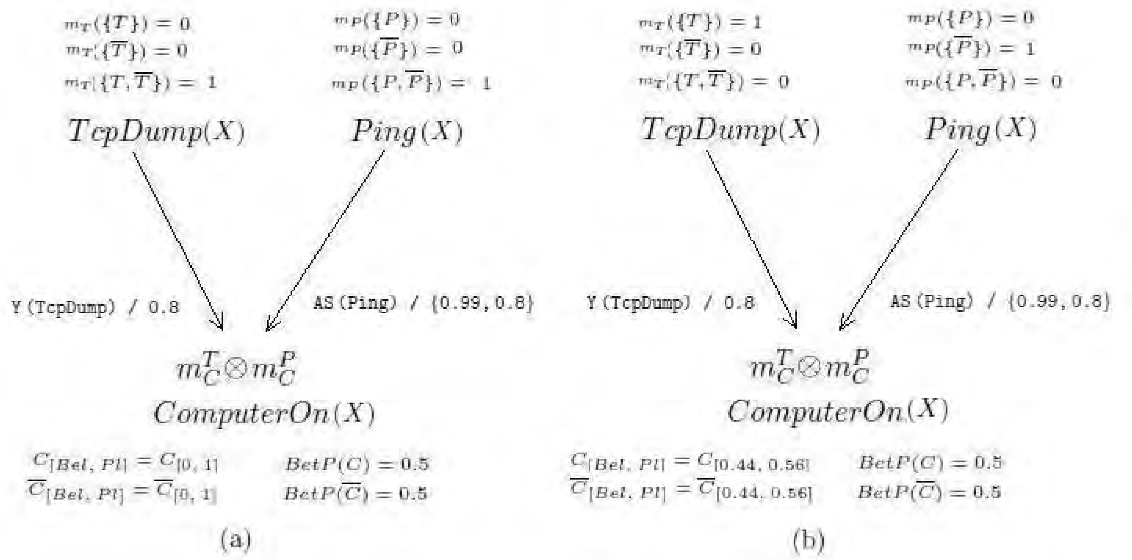


Figure 10: An illustration of disparate situations where the output knowledge provides the same probabilistic answers. In (a), there is absolutely no evidence informing the status of the computer A while in (b) there is an equal amount of evidence that supports both possible states of X. However in both cases, the computer X has an equal probability to be on or off. The propagation functions `Y` *and* `AS` *will be discussed in Section 4.4.*

which allows each proposition A to be interpreted as being either true (i.e., $A_{[1,\,1]}$) , false (i.e., $A_{[0,\,0]}$) or 'unknown' (i.e., $A_{[0,\,1]}$).

4.3 Belief combination

In Section 4.1, we discussed how to combine a set of beliefs pertaining to *different* propositions, $m_{C1}, m_{C2}, \ldots, m_{Cn}$, to form a new belief for the constructed proposition, m_P. For instance, in Figure 6, the *and* operator combines the beliefs associated with $ComputerOn(X)$ and $Classified(X)$ in order to produce the belief for the compound proposition $ClassifiedComputerOn(X)$. This section addresses combination of beliefs pertaining to the *same* proposition — the situation that arises when there is more than one belief, $m_{P1}, m_{P2}, \ldots, m_{Pn}$, induced on P from multiple sources. For example, $ComputerOn(X)$ is at one time associated with both m_C^T and m_C^P propagated from $TcpDump(X)$ and $Ping(X)$ (see Figure 11).

In Influx$_1$, belief combination receives two major treatments: (i) combination by means of a fusion operator (if $m_{P1}, m_{P2}, \ldots, m_{Pn}$ are induced by a set of information sources $\mathbf{E} = \{E_1, E_2, \ldots, E_n\}$ serving as *evidence* for P), and (ii) combination on the basis of sufficiency and necessity criteria (if $m_{P1}, m_{P2}, \ldots, m_{Pn}$ are induced by set of information sources $\mathbf{C} = \{C_1, C_2, \ldots, C_n\}$ serving as *conditions* for P).

The former type of belief combination is extensively studied in the literature of information fusion and evidential reasoning. In these paradigms, E_i serves as *evidence* for P — the evidence E_i may reveal the truth or otherwise of P. With respect to causal knowledge, E_i

corresponds to an *outcome*, *effect* or *manifestation* of P. For instance, '*TcpDump* observes packets from X' (T) is an outcome (thus, evidence) of 'the computer X is *on*' (C) which may in turn be an outcome/evidence for 'the computer X is plugged in' (Pg). As such, by observing T or C, one can infer C or Pg, respectively. When there are multiple pieces of evidence bearing on a proposition (as shown in Figure 11c), the overall belief for the proposition is obtained through an aggregation of the beliefs induced by the evidence using a fusion operator. A fusion operator generally does not assume any form of interaction between the opinions. It aims to find the *consensus* among the opinions in order to obtain a better estimate about the situation.

The latter type of belief combination handles the situations where C_i serves as a *condition* for P — the condition C_i allows the prediction of P. With respect to causal knowledge, C_i corresponds to a *cause* of P. For instance, 'the computer X is plugged in' (Pg) may cause 'the computer X is *on*' (C) which in turn may cause '*TcpDump* observes packets from X'(T). As such, given Pg or C, one may predict C or T, respectively. When there are multiple conditions associated with P, their opinions about the proposition need to be combined according to the notion of sufficiency and necessity to predict P.

When the multiple beliefs associated with a proposition P are induced by both $\mathbf{C}$ and $\mathbf{E}$, the combined belief for $\mathbf{C}$ and the combined belief for $\mathbf{E}$ should be separately computed before being aggregated using a fusion operator. The two types of belief combination are discussed in more detail below.

4.3.1 Belief combination pertaining to evidence

Influx$_1$ is expected to deal with uncertain, incomplete and conflicting information. Such imperfect information can be due to the information itself, to the reliability of its sources, or to errors in measurement, transmission, communication or interpretation of the information. When precision and reliability are critical (such as in situational awareness and mission control systems), information is often collected and fused from multiple sources in order to enhance the accuracy of the information, and subsequently the quality of reasoning and decision making processes.

This type of belief combination is extensively studied in the literature of information fusion and evidential reasoning. Mathematically in the D-S framework, belief combination involves aggregating beliefs induced by multiple information sources bearing on the same frame of discernment (i.e., m_C^T and m_C^P in Figure 11). Here, each information source E_i (or more precisely, the belief of E_i) serves as *evidence* for P, and the beliefs induced by a collection of evidence $\mathbf{E}$ on P are combined with a fusion operator.

In the particular example in Figure 11, $TcpDump(X)$ (T) and $Ping(X)$ (P) serve as information sources for $ComputerOn(X)$ (C), and their respective beliefs m_T and m_P (e.g., '$TcpDump(X)$ believes strongly to have observed packets from X' or '$Ping(X)$ certainly has received responses from X') serve as evidence to infer the status of C. Each individual opinion, m_C^T and m_C^P, corresponds to distinct beliefs bearing on C, and can be viewed as *opinions* of $TcpDump(X)$ and $Ping(X)$ about $ComputerOn(X)$ given m_T and m_P. Belief combination using a fusion operator is usually called for in the following scenarios.

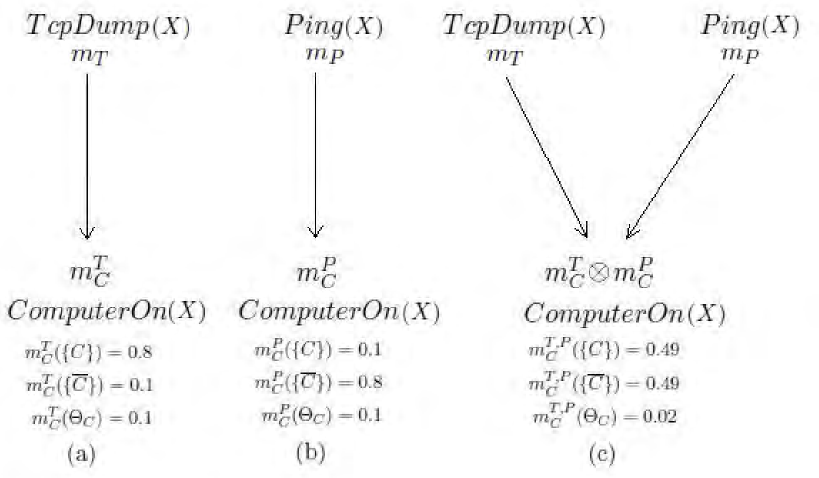


Figure 11: The belief associated with $ComputerOn(X)$ is likely to be more reliable through a combination of opinions (expressed in the form of belief functions m_C^T and m_C^P) provided by the two sensors $TcpDump$ and $Ping$.

- When dealing with imperfect knowledge/information, one often does not rely on one source of information. Rather one attempts to increase the reliability of the information by collecting relevant evidence from multiple sources and combining their respective beliefs about the information in question. Considering the two opinions m_C^T and m_C^P provided by the two sensors $TcpDump$ and $Ping$ given in Figure 11a and Figure 11b, one would obtain a more 'informed opinion' about $ComputerOn(X)$ through an aggregation of the opinions provided by both the sensors (see Figure 11c). As depicted in the figure, m_C^T strongly supports $ComputerOn(X)$ while m_C^P strongly supports its negation $\overline{ComputerOn(X)}$. By fusing these two belief functions using the D-S rule, the obtained $m_C^{T,P}$ reflects the reality that both the proposition and its negation are equally supported by the collected evidence. When an information source is totally ignorant about the situation, its belief (which is expressed in the form of a vacuous bpa) has no impact on the aggregated belief because such a vacuous bpa (m^V) serves as a neutral element, i.e., $m^V \otimes m = m$.

- It is possible that different reasoning paths in a rule network lead to different conclusions regarding the same proposition (e.g., executing the two rules $A \wedge B \rightarrow D$ and $C \rightarrow D$ can result in different states of belief about D). While different conclusions derived for a proposition would be treated as contradictions and thus rejected in monotonic reasoning, this is a typical, expected and desirable situation in Influx$_1$ and the conclusions in such a case are simply aggregated through a combination of their respective belief functions using a fusion operator.

4.3.1.1 Fusion operators A fusion operator is generally used when a more 'informed' view or a more objective opinion about a situation is required through an aggregation of the beliefs pertaining to the different views (or opinions) about the situation. In general, belief fusion promotes belief confirmation/disconfirmation about a proposition A when the different information sources (or evidence) agree with each other in their view about A. More particularly, (i) when the information sources tend to agree with each other in confirming A, belief of A will be promoted, that is a number of weak confirming beliefs might produce a strong confirming belief; and (ii) when the information sources tend to agree with each other in disconfirming A, belief of $\overline{A}$ will be promoted, that is a number of weak disconfirming beliefs might produce a strong disconfirming belief. Conversely, when the information sources are in disagreement, belief fusion appropriately 'balances-out' their opinions (e.g., the D-S rule) or promotes ignorance (e.g., the D-P rule) when computing belief associated with the aggregated opinion.

Having presented such general characteristics of belief fusion, which method to use to perform such aggregation would depend on a number of factors, one being the potential dependency among the opinions (or views) being combined. For instance, let's assume that both a person A and a person B give their opinions that the computer X is on, and let's imagine the two following scenarios: (i) B has actually learnt about this event directly/indirectly from A, and (b) B himself/herself has witnessed the event. It is obvious that a fusion operator should not treat the two opinions as having an equal and independent contribution and thus strengthen the confirmation/disconfirmation of $ComputerOn(X)$ in the former case, while it should in the latter case. This factor, amongst others, explains the existence of a collection of different combination rules investigated in the literature which offer alternative methods to aggregate opinions according to specific situations when relevant knowledge about the opinions is available.

A collection of implemented fusion operators As no single fusion operator investigated in the literature is suitable for all situations, and as Influx_1 is intended to support a wide range of applications, it is necessary for practitioners to have at their disposal a set of operators covering different fusion/reasoning scenarios. To this end, we have implemented the collection of combination rules summarised in Tables 5 and 6, the representatives of which are presented in Table 9.

Due to the adoption of BFoDs as a building block for concept and belief representation in Influx_1, we also include the implementation of a so-called *summation* rule which is applicable when the information sources E_i correspond to mutually exclusive propositions. The summation rule, when applied to combine the two bpas m_1 and m_2 associated with a proposition P, simply sums over all the confirmation $(m_{12}(\{P\}) = m_1(\{P\}) + m_2(\{P\}))$ and disconfirmation belief masses $(m_{12}(\overline{\{P\}}) = m_1(\overline{\{P\}}) + m_2(\overline{\{P\}}))$.

Furthermore, the averaging rule given in Table 5 (Section 3) does not treat vacuous bpas (the bpas that represent total ignorance) as neutral elements as other fusion operators do. Since it is important that a piece of evidence in a total ignorance state should not have impact on the confirmation/disconfirmation of a proposition, we implemented a customised averaging rule in Influx_1 which can effectively and appropriately deal with vacuous belief functions:

$$\frac{m_1(\{A\})(1 - m_1(\Theta_A)) + m_2(\{A\})(1 - m_2(\Theta_A))}{(1 - m_1(\Theta_A)) + (1 - m_2(\Theta_A))},\ A \subset \Theta. \tag{29}$$

From the computational perspective, a bpa m associated with a proposition in Influx_1

is defined only by three numbers, i.e., $m(A)$, $m(\overline{A})$ and $m(\Theta_A)$. This allows simplified versions of the different combination rules to be derived, significantly reducing the computational complexity of the combination process. More specifically, in the case of the D-S rule, a reduced form of this rule is given as

$$(a\,b) \otimes (c\,d) = \left[1 - \frac{\bar{a}\bar{c}}{1-(ad+bc)}\,,\, 1 - \frac{\bar{b}\bar{d}}{1-(ad+bc)}\right] \tag{30}$$

where $\bar{x} = 1-x$, and x, y in $(x\,y)$ denote $Bel(A)$, $Bel(\bar{A})$ of a proposition A, respectively. This particular representation of the rule makes it possible to formulate an inverse function [18] to subtract evidence:

$$(a\,b) - (c\,d) = \left[\frac{\bar{c}(a\bar{d} - \bar{b}c)}{\bar{c}\bar{d} - \bar{b}c\bar{c} - \bar{a}d\bar{d}}\,,\, \frac{\bar{d}(b\bar{c} - \bar{a}d)}{\bar{c}\bar{d} - \bar{b}c\bar{c} - \bar{a}d\bar{d}}\right] \tag{31}$$

Not only does this permit a significant speed-up in combining a large number of belief functions (evidence), but the presence of the inverse formula[15] allows the belief for a concept to be efficiently updated (recomputed by subtracting the belief of the evidence to be removed, and then re-fusing with the belief of the new evidence) without having to repeat the whole combination process[16].

Provided below is a discussion regarding the potential usage of a number of combination rules implemented in Influx$_1$. In particular, it provides some suggestions on the applicability of combination rules in Table 9 as the fusion operator in the various scenarios depicted in Figure 12.

Combination of belief from independent sources In the typical situations depicted in Figures 12a and 12b, the beliefs propagated to D ($A \to D$, $B \to D$, $C \to D$) come from *independent* sources, I-typed operators are the most appropriate ones to combine belief for D. Provided that the sources are reliable (Figure 12a), the D-S and D-P rules are the recommended operators. The D-S rule emphasises consensus between m_D^A, m_D^B, m_D^C and proportionally promotes the states of D in agreement. On the other hand, the D-P rule emphasises both consensus and conflict — it captures 'hard' consensus between opinions and distributes conflict mass to the relevant disjunctive state (or ignorance, in the case of Influx$_1$ — thus equivalent to Yager's rule). For instance, an application of the D-S and D-P rules to the example in Figure 11 would produce the following combined opinions (m_C^{DS} and m_C^{DP}, respectively) about $ComputerOn(X)$:

$$\begin{aligned} m_C^{DS}(\{C\}) &= 0.49, \\ m_C^{DS}(\{\overline{C}\}) &= 0.49, \\ m_C^{DS}(\Theta_C) &= 0.02; \end{aligned}$$

[15] Indeed, an inverse formula exists for the class of non-dogmatic bpas — thanks to Dr. Martin Oxenham for pointing this out to us.

[16] The inverse formula for the D-S rule is not currently implemented in Influx$_1$ due to the fact that there does not yet exist such a formula for the other combination rules. However, the inverse formula can be added to a future version of Influx$_1$ as an optimisation feature, if desired.

Evidence type	Operator type	Specific operator	
Evidence from independent sources	I	I^{DS}	(D-S rule)
		I^{DP}	(D-P rule)
		I^{D}	(disjunctive rule)
Evidence from dependent or partially dependent sources	D	D^{C}	(cautious rule)
		D^{CS}	(semi-cautious rule)
		D^{A}	(averaging rule)
Independent evidence from a common source	C		(cumulative rule)
Evidence from mutually exclusive sources		D^{C}	(cautious rule)
		S	(summation rule)

Table 9: Major fusion operators implemented in Influx$_1$ and their notations used in this document; where I, D, C and S stand for independent, dependent, cumulative and summation, respectively. Please note that as given in Table 5, not all of the fusion operators illustrated here are both commutative and associative (i.e., the averaging rule is not associative and the DP-rule is quasi-associative). However, associativity for these rules is achieved in the implementation of Influx$_1$ through different treatment of the non-associative/associative components of the rules.

and

$$m_C^{DP}(\{C\}) = 0.17,$$
$$m_C^{DP}(\{\overline{C}\}) = 0.17,$$
$$m_C^{DP}(\Theta_C) = 0.66.$$

This example suggests that the use of the D-S rule is recommended in a general case where the information sources are somewhat expected to provide diverging opinions. However, when such divergence may suggest some abnormality or suspicious event that needs to be captured, the D-P rule may become more useful. When there are no conflicts among the opinions, the two rules produce exactly the same results[17].

In some extreme scenarios (Figure 12b) where some of the sensors are known (or assumed) *not* to provide correct information (e.g., due to sensors being broken) but *knowledge of the specific unreliable sources is unavailable*, the more appropriate fusion operator to use in this case would be the disjunctive rule (I^{D}). The disjunctive rule of combination resolves the problem of combining opinions (e.g., m_D^A, m_D^B and m_D^C) potentially provided by unreliable sources by distributing the belief masses associated with partial agreement

[17] Please note that in both cases the degree of conflict associated with the belief functions being combined can be captured in the K value presented in Section 3.

and conflict between the opinions to the relevant disjunctive states of the proposition D (or ignorance, in the case of Influx_1). In general the disjunctive rule operates in a very cautious manner (promoting ignorance), and should only be used when necessary — when most of the sensors are well known to be possibly unreliable.

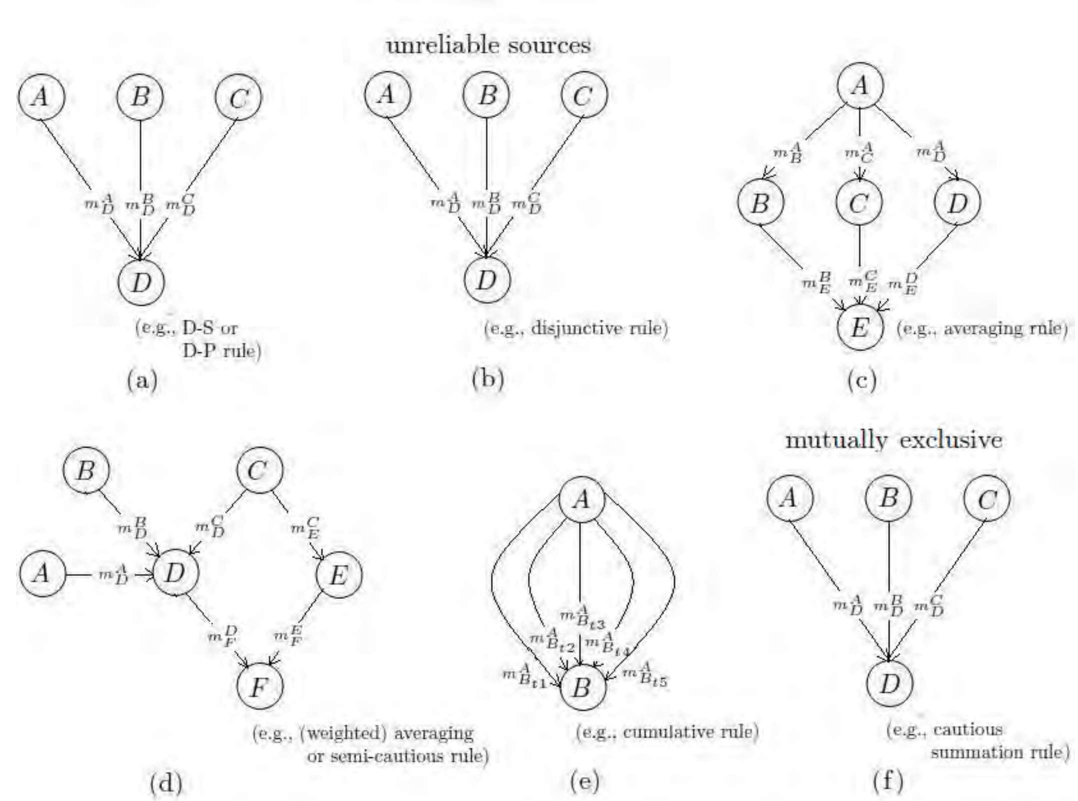


Figure 12: An illustration of various scenarios pertaining to belief combination. In this figure, the belief associated with the information sources for a proposition serve as evidence for the proposition. For instance, the belief associated with A, B and C in (a) are evidence which provide multiple opinions, m_D^A, m_D^B and m_D^C, about D.

Combination of belief from (partly) dependent sources In circumstances where the belief propagated comes from *non-independent* sources (Figures 12c and 12d), D-typed operators can in turn be leveraged.

- With respect to Figure 12c, the beliefs to be combined at E can all be considered to be provided by the same source A, but each 'relayed' through, and adjusted by B, C and D. Thus it is plausible in this case to 'average out' the belief propagated to E from B, C and D, justifying the use of the averaging rule.
- Regarding Figure 12d, the direct information sources of F assume a form of *partial* dependency (i.e., the belief associated with E is influenced by only C while that associated with D is influenced by A, B and C), the applicability of the averaging

rule presented above (which assumes an equal credibility of m_F^D and m_F^E) seems less justified in this case.

- **Customising the averaging rule** One possible solution is to customise the averaging rule and include in its formula the so-called *importance weights* μ_i which reflect the relative importance of each belief function to be combined m_i based on the degree of dependency between their sources. For instance, given the importance weights for m_F^D and m_F^E as μ_1 and μ_2, μ_1 and μ_2 can be computed as 1 and $\frac{1}{3}$, respectively, according to the number of sources feeding evidence to D and E. Then $m_F^D \otimes m_F^E$ can be combined through an incorporation of μ_1 and μ_2 into the averaging rule in Eq. (29) as follows:

$$\frac{\mu_1 m_F^D(\{A\})(1 - m_F^D(\Theta_A)) + \mu_2 m_F^E(\{A\})(1 - m_F^E(\Theta_A))}{\mu_1(1 - m_F^D(\Theta_A)) + \mu_2(1 - m_F^E(\Theta_A))}, \; A \subset \Theta. \tag{32}$$

- **Utilising the cautious rule** Another possible option is to utilise the cautious rule. In some specific scenarios, the cautious rule is able to deal with non-distinct and non-independent evidence by avoiding counting each body of evidence more than once, due to its idempotency. In a general setting, the rule manages the evidence dependency problem by operating in a 'cautious' manner. For instance, the rule first decomposes m_F^D and m_F^E into a collection of simple bpas, $\{m_F^s(\{F\})\}$ and $\{m_F^s(\overline{\{F\}})\}$, and then recombines the two simple bpas that provide the maximal support for $\{F\}$ or $\{\overline{F}\}$. However, the rule assumes a strong evidence dependency and at times can be too 'cautious'. The adequacy of the cautious rule is somewhat diminished when the opinions or pieces of evidence have overlapping information sources, and are thus only in a partial dependency relationship. To this end, Influx$_1$ also offers a so-called *semi-cautious* rule (presented below) to deal with such situations. The usefulness of the semi-cautious rule prevails in cases where the network structure is complex, rendering the measurement of evidence dependency nontrivial or infeasible.

  The semi-cautious rule

  As previously described, the cautious and D-S rules represent both extremes of information fusion with respect to the dependence of evidence. On the one hand, a cautious rule is able to fuse dependent/identical evidence. On the other hand, the D-S rule mandates that the evidence be distinct and independent. To fuse partially dependent evidence, one would need to find an intermediate between these extremes. The idea underpinning the semi-cautious rule is based on the study by Denœux [8] who proposed to replace the minimum operator in the cautious rule's formula (see Table 6) with a parameterised family of triangular-norms (or t-norms)[18] on [0, 1]. One such family of t-norms is the Dubois-Prade family defined as:

$$xT_\gamma^{DP}y = \frac{xy}{max(x, y, \gamma)}, \text{ where T is a a t-norm and } \gamma \in [0, 1] \tag{33}$$

[18] A t-norm is a generalisation of set intersection. Mathematically, a t-norm can be defined as a function from $[0, 1] \times [0, 1]$ to $[0, 1]$ with the following fundamental properties: commutativity, monotonicity, associativity and neutrality of 1.

which include the product and minimum t-norms as special members. When T is the minimum operator (i.e., $\gamma = 0$), one obtains the cautious rule. Conversely, when T is the product (i.e., $\gamma = 1$), one has the D-S rule. In Influx$_1$, the value for γ is encoded as

$$\gamma = max(x, y) + (1 - max(x, y)) * 0.5, \tag{34}$$

thereby achieving the semi-cautious rule. The fusion result of this rule is intermediate between that of the cautious and D-S rules, therefore appropriate for combining beliefs from partly dependent sources.

Combination of belief pertaining to independent pieces of evidence from a common source The cumulative operator should be used when the beliefs being combined are induced by distinct and independent evidence from a common source. As illustrated in Figure 12e, the sensor A is monitoring and regularly reporting information pertaining to B in a temporal manner. For example, $TcpDump$ reports regularly at each time point t_i within a time window t whether it has observed packets generated from the computer X. Since each observation at A is carried out at the time t_i, it is distinct and independent from each other, *yet* provided by the same source A. It is intuitive and plausible in this particular case to compute the belief associated with B for a specific time window t by *accumulating* all the opinions from A to B during the time window (i.e., $m^A_{B_{t1}}$, $m^A_{B_{t2}}$, $m^A_{B_{t3}}$, $m^A_{B_{t4}}$ and $m^A_{B_{t5}}$), justifying the use of the cumulative rule in this case.

Combination of belief from mutually exclusive sources As discussed in Section 4.1.1.3, the BFoDs associated with mutually exclusive propositions can be viewed as coarsenings of an implicit frame Ω. For instance, assuming that A, B and C in Figure 12f constitute an exhaustive set, their respective BFoDs, $\Theta_A = \{A, \overline{A}\}$, $\Theta_B = \{B, \overline{B}\}$ and $\Theta_C = \{C, \overline{C}\}$, are considered coarsenings of $\Omega = \{A, B, C\}$. Specifically, A, B and C correspond to elements (or singleton subsets) of Ω, and $\overline{A}$, $\overline{B}$ and $\overline{C}$ correspond to $\{B, C\}$, $\{A, C\}$ and $\{A, B\}$, respectively. Combination of beliefs associated with D can be treated in the two following ways, which one is to be used depends on the relationship defined between each of Θ_A, Θ_B, Θ_C and Θ_D.

- Provided that the relationship between an information source (e.g., A) and D is defined for all elements of Θ_A (i.e., A and $\overline{A}$), then the opinion propagated from A to D (m^A_D) can be considered an approximation of the opinion to D from its refined frame Ω (m^Ω_D). To this end, m^A_D, m^B_D and m^C_D would correspond to variations of the same opinion m^Ω_D. Thus it is plausible to combine the opinions using the cautious rule which treats the opinions as carrying overlapping/duplicate information in the combination process.

- Alternatively, the relationship between an information source (e.g., A) and D may be only defined for A (i.e, given $\overline{A}$, it is ignorant about D). As such, the opinions m^A_D, m^B_D and m^C_D do not represent variations of the opinion from Ω, but correspond to the 'elementary' opinions induced by individual elements of Ω. In this case, m^A_D, m^B_D and m^C_D have an additive effect on D and thus should be combined using the summation rule.

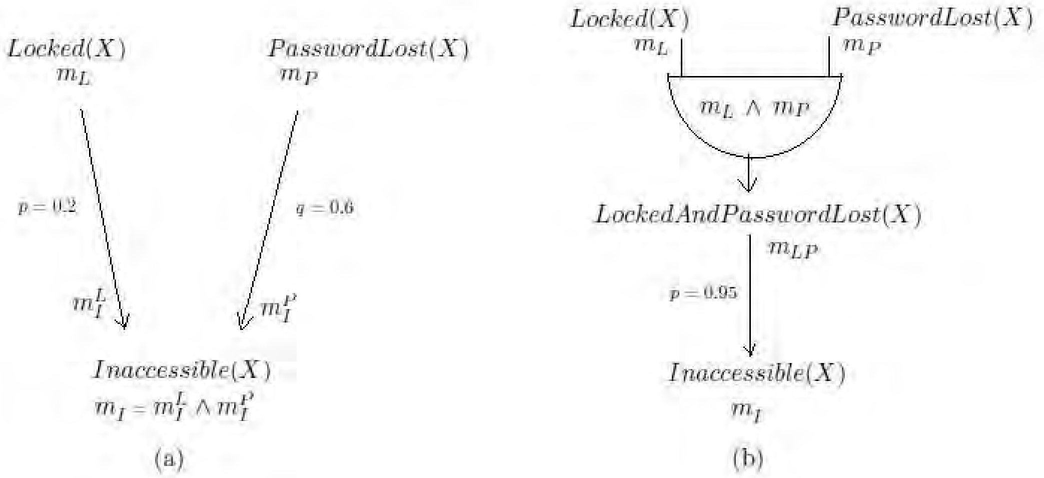


Figure 13: An illustration that distinguishes the use of the 'and' operator to (i) combine opinions propagated to $Inaccessible(X)$ from different information sources (i.e., $m_I^L \wedge m_I^P$ as in (a)) with the use of the operator to (ii) combine the information sources themselves (i.e., $m_L \wedge m_P$ as in (b)) before propagating belief to $Inaccessible(X)$. The latter situation is appropriate and valid where the conditions $Locked(X)$ and $PasswordLost(X)$ have a synergistic effect on the proposition $Inaccessible(X)$.

4.3.2 Belief combination pertaining to conditions

When P may be predicted by of one or more conditions $\mathbf{C} = \{C_1, C_2, \ldots, C_n\}$, the operator that combines the beliefs induced by $\mathbf{C}$ on P should reflect the sufficiency and/or necessity of the conditions with respect to P. To this end, belief combination can be performed using the various operators presented in Table 7 (Section 4.1.1.3) in place of fusion operators. Practically, such operators can be utilised in one of two ways: (i) constructing a new proposition (or condition) C' from C_i (if elements in $\mathbf{C}$ collectively constitute a condition for P), and (ii) combining the beliefs (or opinions) induced by each C_i on P (if elements in $\mathbf{C}$ serve as individual conditions for P).

Elements in C collectively constitute a condition for P The influence on P from $\mathbf{C}$ in this case is captured in a single relation. This requires a new proposition C' constructed from C_i as discussed in Section 4.1.1.3, and a specified relationship between C' and P. Examples are shown in Figures 13b and 14b. This approach is generally applicable in the following scenarios.

- The truth of all the conditions C_i in $\mathbf{C}$ has a synergistic effect on P (see Figure 13b). In this case, it is necessary to construct a new proposition C' that reflects the conjunctive truth of C_i using an *and*-typed operator. For instance, each of the propositions $Locked(X)$ and $PasswordLost(X)$ on its own may not have a very strong effect on whether the computer X is currently inaccessible (see Figure 13a). However, the fact that both of the conditions are true has a synergistic effect on

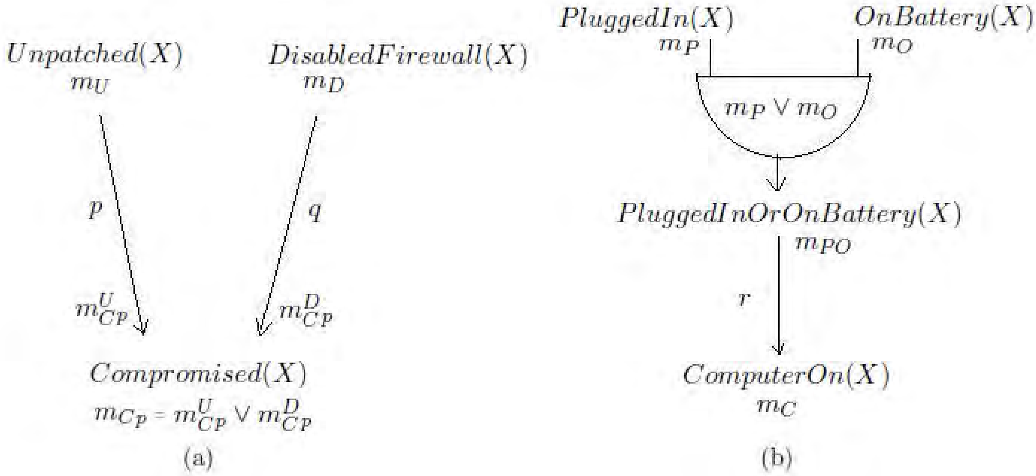


Figure 14: An illustration that distinguishes the use of the 'or' operator to i) combine opinions propagated to $Compromised(X)$ from different information sources (i.e., $m^U_{Cp} \vee m^D_{Cp}$ as in (a)) with the use of the operator to (ii) combine the information sources themselves (i.e., $m_P \vee m_O$ as in (b)) before propagating belief to $ComputerOn(X)$. The latter approach is appropriate when the relationship between $ComputerOn(X)$ and its conditions $PluggedIn(X)$ and $OnBattery(X)$ is not sensitive to the specific condition, or conditions, that are true.

the inaccessibility of X, in which case it is necessary and important to express the influence among the propositions as presented in Figure 13b. Expressing the influence from $\mathbf{C}$ onto P as depicted in Figure 13a would not reflect the reality, thereby providing an incorrect answer.

- The truth of one or more conditions C_i has an impact on P, however the nature of the relation between $\mathbf{C}$ and P is not sensitive to the specific condition C_i that is true, nor which combination of C_i that is true (see Figure 14b). In this case, an *or*-typed operator (if C_i serve as alternative conditions for P), or the *mean* operator (if C_i are neither necessary nor sufficient) can be used to construct the new proposition C'. For instance, $ComputerOn(X)$ can be considered confirmed if one or more of the conditions ($PluggedIn(X)$ and $OnBattery(X)$) is confirmed, without a need to clearly distinguish which condition, or which group of conditions, is confirmed. As such the relationship between the propositions can be modelled as presented in Figure 14b.

  When the specific knowledge pertaining to the strength of the relation between each element of $\mathbf{C}$ and P is unavailable, this approach can also be used to approximate the belief computed for P. Practical advantages of this approach include the provision of a simpler rule network, fewer numerical values required as input, and an alleviation of the complexity associated with knowledge acquisition, learning and training.

Elements in C serve as individual conditions for P Each condition C_i may have a different impact on P, and the impact on P may be increased with the number of C_i that is confirmed. This requires a specified relationship between each C_i and P, and combination of the beliefs (or opinions) induced by C_i on P as shown in Figure 14a. In the figure, $Compromised(X)$ is confirmed to a degree of p and q if either of the conditions $Unpatched(X)$ and $DisabledFirewall(X)$ is confirmed, respectively. The support for $Compromised(X)$ is only increased (linearly to p and q) if both $Unpatched(X)$ and $DisabledFirewall(X)$ are confirmed. This necessitates the use of the *or* operator to combine the belief induced on p from the conditions (as shown in Figure 14a). Similarly, other operators in Table 7 can be used to combine the opinions from C_i according to the dependence/independence between C_i, and the sufficiency/necessity between C_i and P. Exceptions are and/and^{sub} which are generally applicable when elements in $\mathbf{C}$ collectively constitute a condition, and rarely find application in the type of combination discussed in this approach.

4.4 Belief propagation

The previous section deals with aggregating belief functions defined on the same frame of discernment induced by different sources (e.g., m_C^T and m_C^P in Figure 15) for which a collection of combination methods implemented in Influx_1 have been presented and discussed. This section is concerned with belief transfer between different frames of discernment, such as belief transfer between m_T and m_C^T, or between m_P and m_C^P. Here, the context is that while a portion of propositions in the knowledge base receive evidence from dedicated sources or sensors, the majority of them would receive no direct observations, and thus assessment of their belief must be inferred (if possible) from other propositions. To this end, belief propagation allows one to infer the degree of belief for those propositions which the collected evidence do not directly bear on.

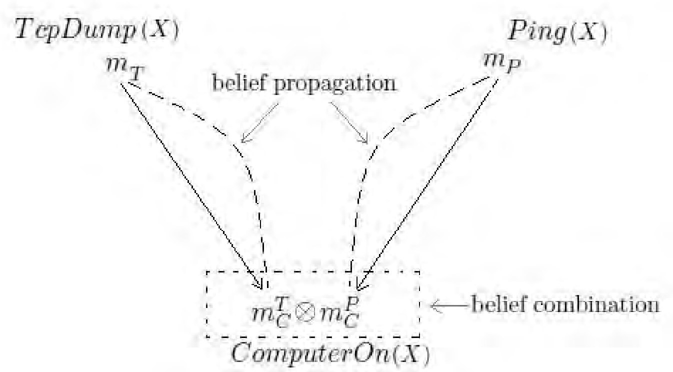


Figure 15: An illustration of belief combination and belief propagation.

Given any two propositions with a dependency relationship, such a relation can be implemented as an *implication* (or '*if-then*') rule which is also associated with a belief quantifying the strength of the relationship:

$$A \xrightarrow{p} B, \text{ where } p \in [0,1]. \tag{35}$$

For instance, the relation between $TcpDump(X)$ and $ComputerOn(X)$ might be described as:

$$TcpDump(X) \overset{0.8}{\rightarrow} ComputerOn(X). \qquad (36)$$

In the Bayesian probabilistic interpretation, p is understood as a probabilistic measure, and Eq. (35) is the conditional probability of B given A. As such, Eq. (36) can be read as: given that TcpDump observes packets from X, there is an 80% chance that X is on, directly entailing a 20% chance that X is off.

Unlike the probabilistic standpoint, D-S reasoning is focused on providing 'provable' belief measuring the truth of hypotheses/propositions based on the available evidence. To this end, the quantity p in Eq. (35) is not a probabilistic measure, but a belief mass[19] indicating the extent to which B is *logically supported* by the evidence A. In this sense, Eq. (36) can no longer automatically infer 0.2 as the degree of support for $\overline{ComputerOn(X)}$ since the fact that TcpDump observes packets from X does *not* serve as a piece of evidence proving that the computer X is off. Given that $TcpDump$ observes packets from X (i.e., $m_T(\{T\}) = 1$, $m_T(\{\overline{T}\}) = 0$ and $m_T(\Theta_T) = 0$), it is thus more appropriate to assign belief masses to $ComputerOn(X)$ (denoted as C) as follows:

$$\begin{aligned} m_C^T(\{C\}) &= 0.8, \\ m_C^T(\{\overline{C}\}) &= 0, \\ m_C^T(\Theta_C) &= m_C^T(\{C, \overline{C}\}) = 0.2, \end{aligned}$$

where 0.8 is the degree of belief for $\{C\}$ supported by the available knowledge and evidence, $\{\overline{C}\}$ is currently not supported by any evidence, and 0.2 is the amount of uncertainty waiting to be transferred to either $\{C\}$ or $\{\overline{C}\}$ in light of new evidence. As such, a more precise presentation of Eq. (35) would be

$$A \overset{[p,\,1]}{\rightarrow} B \qquad (37)$$

where $[p,\ 1]$ corresponds to the belief interval $[Bel,\ Pl]$ associated with the implication rule. For simplicity of notation, Eq. (35) is usually used in place of Eq. (37) throughout the document. In a number of cases, the belief of A may influence the belief of both B and $\overline{B}$. For instance, one may assert that if the the computer X is plugged in, it is believed to be 'on' to a degree of p. However, if the person is also aware that X has loose cables, and thus it may not be on when plugged in to a degree of q. In this case, the relation between the two propositions is given as

$$PluggedIn(X) \overset{[p,\,1-q]}{\rightarrow} ComputerOn(X).$$

When $q = 1 - p$ (i.e., $A \overset{[p,\,p]}{\rightarrow} \overline{B}$), the belief function computed for B will be equivalent to that of the Bayesian implication.

4.4.1 Belief propagation functions

As stated above, Eq. (35) corresponds to the standard logical implication which provides an endorsement for the conclusion based on the belief measuring the truth of the premise.

[19] Such a p value in Influx_1 is also referred to as a *weight* that quantifies the strength of the relation between any two propositions A and B.

More specifically, on observing packets generated from X, the sensor $TcpDump$ can predict with a high degree of confidence (i.e., at least 80%) that a computer X is on. Otherwise, the sensor is not able to provide any information regarding this proposition (i.e., the fact that $TcpDump$ has not seen any packets from X does not reveal information regarding whether the computer is on).

In this formulation, belief can be propagated in a similar way to that of logical rule-based reasoning where the belief of the premise (a simple or compound proposition) is aggregated with the belief associated with the rule itself in order to deduce the belief concerning the conclusion. Our concern here is how to combine the belief associated with the premise with that associated with the implication rule in such a way that it is consistent with D-S theory.

Since combination of belief in the D-S framework is based on set-theoretic operations, it is necessary for the belief functions being combined to be defined on the same frame of discernment (or the same proposition in the case of Influx$_1$). This means, in order to express and perform the influence of belief between A and B in Eq. (35) (each associated with a BFoD, Θ_A and Θ_B, respectively), a new frame needs to be created in such a way that it can distinguish all the propositions in the original frames. Provided that Θ_A and Θ_B are independent, one such frame is the Cartesian product of Θ_A and Θ_B:

$$\Theta_{A,B} = \{(A,B), (A,\overline{B}), (\overline{A},B), (\overline{A},\overline{B})\}$$

This new frame is indeed a common refinement of both Θ_A and Θ_B (more details about frame coarsenings and refinements can be found in Section 3) on which A and B can be both expressed using the following mapping rules:

$$\begin{aligned} \omega_1 &: 2^{\Theta_A} \to 2^{\Theta_{AB}}, \\ \omega_1(\{A\}) &= \{(A,B),(A,\overline{B})\}, \\ \omega_1(\overline{\{A\}}) &= \{(\overline{A},B),(\overline{A},\overline{B})\}, \end{aligned} \tag{38}$$

$$\begin{aligned} \omega_2 &: 2^{\Theta_B} \to 2^{\Theta_{AB}}, \\ \omega_2(\{B\}) &= \{(A,B),(\overline{A},B)\}, \\ \omega_2(\overline{\{B\}}) &= \{(A,\overline{B}),(\overline{A},\overline{B})\}. \end{aligned} \tag{39}$$

Then by combining $m_{(A,B)}$ and m_A (which defines the current state of belief for A) using the D-S rule, and subsequently reducing the resulting belief function on Θ_B, one would obtain m_B which defines the belief induced on B (as shown in Example 8).

Example 8 Given the implication rule in Eq. (36), this rule can be interpreted as *not* T *or* C (where T and C denote $TcpDump(X)$ and $ComputerOn(X)$, respectively) and expressed on $\Theta_{(T,C)}$ in the form of the following belief function [51]:

$$\begin{aligned} m^R_{(T,C)}(\{(T,C),(\overline{T},C),(\overline{T},\overline{C})\}) &= 0.8, \\ m^R_{(T,C)}(\Theta_{(T,C)}) &= 0.2. \end{aligned}$$

Let's assume that the sensor $TcpDump$ believes to the degree of at least 90% of certainty that it has observed packets generated from X. Thus the bpa associated with the premise of Eq. (36) is:

$$\begin{aligned} m_T(\{T\}) &= 0.9, \\ m_T(\{\overline{T}\}) &= 0, \\ m_T(\Theta_T) &= 0.1, \end{aligned}$$

which can be defined on $\Theta_{(T,C)}$ using Eq. (39) as

$$\begin{aligned} m^T_{(T,C)}(\{(T,C),(T,\overline{C})\}) &= 0.9, \\ m^T_{(T,C)}(\Theta_{(T,C)}) &= 0.1, \end{aligned}$$

Expressed on the same frame of discernment, $m^R_{(T,C)}$ and $m^T_{(T,C)}$ can now be combined using the D-S rule. The combined belief function that results is

$$\begin{aligned} m^{T,R}_{(T,C)}(\{(T,C)\}) &= 0.72, \\ m^{T,R}_{(T,C)}(\{(T,C),(T,\overline{C})\}) &= 0.18, \\ m^{T,R}_{(T,C)}(\{(T,C),(\overline{T},C),(\overline{T},\overline{C})\}) &= 0.08, \\ m^{T,R}_{(T,C)}(\Theta_{(T,C)}) &= 0.02. \end{aligned}$$

By reducing $m^{T,R}_{(T,C)}$ on Θ_C, one obtains the following belief function for $ComputerOn(X)$:

$$\begin{aligned} m_C(\{C\}) &= 0.72, \\ m_C(\{\overline{C}\}) &= 0.0, \\ m_C(\Theta_C) &= 0.18 + 0.08 + 0.02 = 0.28. \end{aligned}$$

Indeed, by defining a relation between A and B as a belief function on $\Theta_{(A,B)}$, Influx$_1$ provides a mechanism to express an arbitrary relationship pertaining to A and B, rather than being limited to a 'strict' logical implication (as investigated in [9], [17], [15] and [18]).

For the purpose of enhancing the usability of the tool and minimising the computational overhead at run-time, Influx$_1$ reduces the need for users to define a relation between any two propositions A and B as a belief function on $\Theta_{(A,B)}$, by providing a collection of built-in functions that capture a number of potentially useful relations between propositions in practice. These functions allow belief propagation to be directly carried out using the derived formulas, thereby bypassing the whole combination process presented in Example 8.

Table 10 illustrates a subset of the common types of relations between any two propositions and Table 11 exemplifies how the relations are formulated and implemented in Influx$_1$. In the case where the known relationship between any two propositions A and B is more complex, and thus cannot directly be captured using the pre-defined functions (such as those presented in Table 11), the relation between A and B could be expressed in Influx$_1$ using a generic assignment function with additional parameters:

```
B() = AS(A() / {{p₁, p₂}, {q₁, q₂}, {r₁, r₂}})
```

Type	Relation	Relation description	Example	Example description
Single relation	$A \xrightarrow{p} B$	Relation between the two propositions is known	$TcpDump(X) \xrightarrow{0.8} ComputerOn(X)$	If *TcpDump* observes packets originating from X, it is likely that the computer X is on; otherwise it is unknown whether X is on
	If $p = 1$	The above becomes an *is_a* relation	$Ubuntu(X) \xrightarrow{1} Linux(X)$	If X is *Ubuntu*, X is certainly *Linux*.
	$\overline{A} \xrightarrow{q} \overline{B}$	Relation between the negation of the two propositions is known	$\overline{pluggedIn(X)} \xrightarrow{1} \overline{ComputerOn(X)}$	If the computer X is *not* plugged in, it is certainly *not* on; otherwise, it is unknown if X is on
Parallel relation	$A \xrightarrow{p} B$, $\overline{A} \xrightarrow{q} \overline{B}$	Relations between the two propositions, and between their negation, are both known	$Ping(X) \xrightarrow{0.99} Computer(X)$, $\overline{Ping(X)} \xrightarrow{0.8} \overline{Computer(X)}$	If *Ping* receives a response from the computer X, it is almost certain that X is on; if *Ping* does not, it is likely that X is not on
	If $p = 1$, $q = 1$	The above becomes an *equivalent* relation	$ComputerOn(X) \xrightarrow{1} PowerUsed(X)$, $\overline{ComputerOn(X)} \xrightarrow{1} \overline{PowerUsed(X)}$	If the computer X is on, it consumes power; if it is off, it does not consume power
	$A \xrightarrow{p} \overline{B}$, $\overline{A} \xrightarrow{q} B$	Relations between one proposition and the negation of the other, as well as between the negation of the proposition and the other proposition, are both known	$Ping(X) \xrightarrow{0.4} \overline{Firewalled(X)}$, $\overline{Ping(X)} \xrightarrow{0.7} Firewalled(X)$	If *Ping* receives a response from an active computer X, it is somewhat likely that X does not have an active firewall; if *Ping* does not, it is likely that X has an active firewall
	If $p = 1$, $q = 1$	The above becomes a *not* relation	$Classified(X) \xrightarrow{1} \overline{Unclassified(X)}$, $\overline{Classified(X)} \xrightarrow{1} Unclassified(X)$	If X is classified, it is certainly not unclassified and vice versa
	If $p = 1$, $q = 0$,	The above becomes an *is_not_a* relation	$Ubuntu(X) \xrightarrow{1} \overline{RedHat(X)}$, $\overline{Ubuntu(X)} \xrightarrow{0} RedHat(X)$	If X is *Ubuntu*, it is certainly not *RedHat*; otherwise, it is unknown if X is *RedHat*

Table 10: A subset of common types of relation between any two propositions.

Type	Relation	Relation expression on $\Theta_{(A,B)}$	Direct formula	Relation expression in Influx$_1$
Single relation	$A \xrightarrow{p} B$	$m_{A,B}(\{(A,B),(\overline{A},B),(\overline{A},\overline{B})\}) = p$ $m_{A,I}(\Theta_{(A,B)}) = 1-p$	$m_B(B) = p * m_A(A)$ $m_B(\overline{B}) = 0$ $m_B(\Theta_B) = 1-p*m_A(A)$	`B()=Y(A()/p)` (Yes transform)
	If $p = 1$		$m_B(B) = m_A(A)$ $m_B(\overline{B}) = 0$ $m_B(\Theta_B) = 1 - m_A(A)$	`B()=Y(A())` `B()=INC(A())`
	$\overline{A} \xrightarrow{q} \overline{B}$	$m_{A,B}(\{(\overline{A},\overline{B}),(A,B),(A,\overline{B})\}) = q$ $m_{A,I}(\Theta_{(A,B)}) = 1-q$	$m_B(B) = 0$ $m_B(\overline{B}) = q * m_A(\overline{A})$ $m_B(\Theta_B) = 1-q*m_A(\overline{A})$	`B()=N(A()/q)` (No transform)
Parallel relation	$A \xrightarrow{p} B$, $\overline{A} \xrightarrow{q} \overline{B}$	$m_{A,B}(\{(A,B),(\overline{A},B),(\overline{A},\overline{B})\}) = p$ $m_{A,B}(\Theta_{(A,B)}) = 1-p$ $m'_{A,B}(\{(\overline{A},\overline{B}),(A,B),(A,\overline{B})\}) = q$ $m'_{A,B}(\Theta_{(A,B)}) = 1-q$	$m_B(B) = p * m_A(A)$ $m_B(\overline{B}) = q * m_A(\overline{A})$ $m_B(\Theta_B) = 1-p*m_A(A)$ $-q * m_A(\overline{A})$	`B()=AS(A()/{p,q})` (Assignment)
	If $p = q$		$m_B(B) = p * m_A(A)$ $m_B(\overline{B}) = p * m_A(\overline{A})$ $m_B(\Theta_B) = 1-p*m_A(A)$	`B() = AS(A()/p)`
	$A \xrightarrow{p} \overline{B}$ $\overline{A} \xrightarrow{q} B$	$m_{A,B}(\{(A,\overline{B}),(\overline{A},B),(\overline{A},\overline{B})\}) = p$ $m_{A,B}(\Theta_{(A,B)}) = 1-p$ $m'_{A,B}(\{(\overline{A},B),(A,B),(A,\overline{B})\}) = q$ $m'_{A,B}(\Theta_{(A,B)}) = 1-q$	$m_B(\overline{B}) = p * m_A(A)$ $m_B(B) = q * m_A(\overline{A})$ $m_B(\Theta_B) = 1-p*m_A(A)$ $-q * m_A(\overline{A})$	`B()=NOT(A()/{p,q})` (Not transform)
	If $p = 1$, $q = 1$		$m_B(B) = m_A(\overline{A})$ $m_B(\overline{B}) = m_A(A)$ $m_B(\Theta_B) = m_A(\Theta_A)$	`B()=NOT(A())`
	If $p = 1$, $q = 0$		$m_B(\overline{B}) = m_A(A)$ $m_B(B) = 0$ $m_B(\Theta_B) = 1 - m_A(A)$	`B()=NOT(A()/{1,0})` `B()=NOT^INC(A())`

Table 11: An illustration of how the relations presented in Table 10 are formulated and implemented in Influx$_1$. Given in the first and second columns are the types of relation between the two propositions A and B; given in the third column is the representation of the relations in the form of belief functions on $\Theta_{(A,B)}$; detailed in the fourth column are the direct formulas to compute belief for B derived from the combination process presented in Example 8, and presented in the last column are the corresponding expressions in Influx$_1$.

which defines a wide range of relations between A and B:

$$A \overset{[p_1\,,\,1-p_2]}{\rightarrow} B,$$

$$\overline{A} \overset{[q_1\,,\,1-q_2]}{\rightarrow} B,$$

$$\{A, \overline{A}\} \overset{[r_1\,,\,1-r_2]}{\rightarrow} B.$$

Here, p_1 and p_2 indicate the extent to which B is believed to be confirmed, or disconfirmed, by A; q_1 and q_2 the extent to which B is believed to be confirmed, or disconfirmed, by $\overline{A}$, and r_1 and r_2 the extent to which B is believed to be confirmed, or disconfirmed, by $\{A, \overline{A}\}$ (or Θ_A), respectively. This allows the belief to be computed for B using the following formulas:

$$\begin{aligned} m(B) &= p_1 * m(A) + q_1 * m(\overline{A}) + r_1 * m(\Theta_A), \qquad (40)\\ m(\overline{B}) &= p_2 * m(A) + q_2 * m(\overline{A}) + r_2 * m(\Theta_A),\\ m(\Theta_B) &= 1 - m(B) - m(\overline{B}). \end{aligned}$$

4.4.2 Utility functions

Beyond the collection of functions implemented for the specific purpose of proposition construction, belief combination and belief propagation as presented, Influx$_1$ also provides a number of utility functions. These utility functions exist to provide users with ways to manipulate belief masses associated with propositions in ways that may be useful for various applications. For instance, the *conflict function* (to measure the degree of internal conflict within a belief between a proposition and its negation), and the *ignorance function* (to measure the degree of ignorance in a belief function).

4.5 A brief discussion of other reasoning aspects

In this section, we briefly mention other reasoning aspects of Influx$_1$ which are not among the main topics of discussion in the document, including forward and backward chaining, dynamic knowledge and belief change, utilisation of predicate calculus, and interfaces to external programs.

4.5.1 Forward and backward chaining

Forward chaining is a process for propagating belief updates to related propositions (via directed links), instantiating new propositions if possible. Forward chaining occurs whenever the belief for an existing proposition changes, or a new proposition is instantiated. As an optimisation, only belief changes with a magnitude that exceeds a threshold T are forward propagated. Since the impact of a belief change may diminish through multiple propagations, this can become a significant efficiency gain, especially in large networks.

Conversely, beginning with a proposition being queried, backward chaining is a process for instantiating new propositions as required to derive the query result, following directed links in a backward direction. This may occur when a proposition is queried that is not yet instantiated. When new propositions are instantiated as a result of backward chaining, forward chaining for each such proposition will occur.

4.5.2 Dynamic knowledge and belief change

Influx$_1$ is designed to operate in real-time with dynamic knowledge and beliefs. More specifically, Influx$_1$ supports the following operations during run time:

- belief values for propositions can be changed,
- propositions can be added, or removed, and
- rules (that define proposition construction, belief propagation, or belief combination) can be added.

The above operations may be followed by forward and backward chaining processes to automatically instantiate and update beliefs. Unless an entire network of propositions is removed, the removal of a proposition might be inhibited by the network dependency structure. In such a case, effective removal may instead be achieved by assigning a fixed belief of ignorance to the proposition.

4.5.3 Representing knowledge with predicates and variables

For many purely numerical approaches for reasoning under uncertainty, propositional representations are used. With such a representation, each relation must be explicitly stated for each object in a domain:

if one can ping *1.1.1.1*, it is likely that the computer at *1.1.1.1* is *on*,
if one can ping *1.1.1.2*, it is likely that the computer at *1.1.1.2* is *on*,
if one can ping *1.1.1.3*, it is likely that the computer at *1.1.1.3* is *on*,
etc.

Instead, Influx$_1$ adopts the predicate calculus, where a predicate acts as a template that describes a property or relationship represented by variables (denoted by the prefix \$), allowing a relation to be stated for all objects in a domain:

if one can ping $\$Address$, it is likely that the computer at $\$Address$ is *on*,

$$\text{(i.e., } Ping(\$Address) \rightarrow ComputerOn(\$Address)\text{).}$$

A predicate may be associated with n variables (n-ary predicate). Unary predicates ($n = 1$) are used to describe a property or membership of instances represented by the variable, such as $ComputerOn(\$Address)$, and n-ary predicates ($n > 1$) are used to express relationships among variables, such as $Communicating(\$Sender, \$Receiver)$. The scope of such variables is limited to the specific relation in which they appear, e.g., the variables $\$Sender$ and $\$Receiver$ in the following relation only take on values within the relation:

$$SendingPacket(\$Sender, \$Receiver) \rightarrow Communicating(\$Sender, \$Receiver).$$

To enable reasoning across contexts, Influx_1 supports the use of structured predicate names to embed contextual information. In the example below which states whether the software $MozillaFirefox$ is installed on the computer $fileserver$, $Installed$ carries the semantics defined in the context of $Computer.Software$:

$$Computer.Software.Installed(fileserver, MozillaFirefox).$$

4.5.4 Distributed reasoning

Influx_1 also supports the distribution of reasoning, desirable for large reasoning tasks and applications benefitting from decentralisation. With this capability, a reasoning network can be divided into fragments each stored and executed in one or more Influx instances. The collection of such instances is distributed among a pool of computers, sending and receiving knowledge, and coordinating the results from each other in order to perform a certain global task. Not only does this allow for rapid reasoning and/or reasoning at scale, but it also provides flexibility and redundancy to enhance system resilience and robustness against intentional or accidental incidents.

4.5.5 Interfaces to external programs

Not all types of problems are naturally suited to being entirely modelled and implemented within Influx_1. In general, problems such as those that require complex control knowledge with significant looping, branching and dependent subprocesses, or that call for complicated arithmetic computation, would be difficult to solve entirely within Influx_1. Such problems might be solved at a meta-level by having an external program that interacts with instances of Influx_1. As such, Influx_1 is intended to work in conjunction with external programs, such as other reasoning engines and control programs. To this end, Influx_1 provides a number of interfaces (currently Unix pipes, network sockets, Influx files and the console) that provide common services. This allows an application to be built more naturally, where each tool is used to its particular strength.

5 Final discussion and remarks about Influx$_1$

Influx$_1$ is a simple, but highly flexible and efficient, tool and framework for general reasoning with uncertainty. More specifically, Influx$_1$ is capable of handling imperfect knowledge, operating in a distributed and dynamic manner, while attaining plausible reasoning due to its various methods to represent and deal with uncertainty formulated based on D-S theory. The design and development of Influx$_1$ is driven by the practical objectives stated at the beginning of the document. The following discussions and remarks are with respect to these objectives. Empirical results for the application of Influx$_1$ to certain problems of interest may be found in a forthcoming document.

Flexibility The flexibility of Influx$_1$, due to a large extent to the utilisation of D-S belief representation, fusion and inference, manifests in many aspects:

- Regarding input knowledge, unlike classical reasoning (which treats knowledge with an absolute certainty) and those reasoners based on probability theory (which demands 'completeness' of knowledge), Influx$_1$ is able to represent and reason with uncertain and incomplete knowledge. In other words, users are only required to provide information that is *available*. If a piece of knowledge is known with an absolute certainty, it can be entered into Influx$_1$ as a fact. Conversely, if the piece of knowledge is associated with a specific degree of uncertainty, the uncertainty pertaining to the knowledge can be quantified in the form of objective probability if relevant statistical data exists, or subjective belief if such statistical information is missing but expert knowledge is available. In the cases where statistical information or expert knowledge is available but not complete, one is able to provide such knowledge, or simply state 'total ignorance' in the absence of all the mentioned knowledge or information.

- With respect to knowledge structure, Influx$_1$ allows a wide range of relations between propositions to be expressed. For usability and efficiency purposes, direct formulas for a number of commonly encountered relations between propositions are derived and implemented in Influx$_1$ in the form of built-in functions available for users. Influx$_1$ also supports dynamic knowledge structures, with rules having a variable number of inputs (or information sources).

- In terms of reasoning, the D-S based inference in Influx$_1$ is capable of simplifying into probabilistic reasoning, or reducing to three-valued logical reasoning in a seamless manner at any time during its execution as demanded by the users or warranted by the situations. The reasoning framework of Influx$_1$ can also be directly extended to facilitate inferences in different modalities, e.g., inferencing in a possibilistic manner where observations are specified in a form of 'possibilistic evidence' and the D-S rule replaced by a possibilistic fusion operator, or deductive reasoning with binomial opinions using subjective logic where belief combination is to be carried out using Jøsang's consensus rule.

- With respect to output knowledge, Influx$_1$ provides a rich and informative interpretation of knowledge. It provides information which indicates both *the degree to* *which a proposition/hypothesis is believed to be true* based on the available evidence (by means of the belief interval $[Bel, Pl]$) and *the degree to which the proposition is probably true* (by means of the $BetP$ function). Reasoners based on Bayesian models are usually capable of providing answers only for the latter case.

- Influx_1 also allows developers to capture the degree of conflict among pieces of evidence or knowledge. Not only does Influx_1 offer various ways to resolve conflicts that best suit the specific applications and situations at hand, but it is also the case that the captured degree of conflict itself can potentially play an important role in detecting abnormality and deception. For instance, a deception model could be built within or on top of Influx_1 which utilises application/domain-specific knowledge regarding the information sources and hypotheses (or propositions) of interest, together with the relevant conflicts dynamically captured at run-time in order to detect suspicious events and phenomena.
- Influx_1 provides interfaces with common services, allowing other programs and tools to interact with instances of Influx_1. This allows each portion of a problem to be solved using the most appropriate approach, utilising Influx_1 and other programs where favourable, providing greater flexibility in solving problems with the assistance of Influx_1. These same interfaces also allow multiple instances of Influx_1 to interact with each other, which encourages modular and distributed solutions for further flexibility.

Efficiency The efficiency of Influx_1 is due to the combination of highly-optimised techniques and methods for data storage and algorithms utilised at the implementation level, and the simplicity of the design.

- Regarding the implementation of Influx_1, the achieved high performance is due to its use of Hold (a multiple key hash table implementation with Bloom filters) that enables rapid lookup, search and match operations (for knowledge queries and rules), and the compact storage of knowledge, the details of which are for another document.
- Regarding the high-level design, Influx_1's efficiency is gained through local computations and the utilisation of coarsened (binary) frames of discernment to represent the portion of the discernment space of interest. This allows the derivation of significantly simplified formulas for all DS-related rules and inferences.

  Unlike a number of existing rule-based systems which do not (or to a limited degree) take into account the potential dependency of evidence being combined, Influx_1 capitalises on state-of-the-art fusion methods and techniques in the literature in providing users with mechanisms to more effectively tackle this problem. More specifically, the reasoner offers a range of operators that allow users to combine beliefs induced from independent sources, from dependent sources, and to approximate the combination result when the dependency of evidence is potentially complex beyond what can be captured and expressed in an intuitive and plausible manner.

Scalability Though the use of numerical values for uncertainty representation is convenient and useful, quantitative approaches often raise practical concerns about *the potentially large number of numerical values that are involved and their required precision*, which can hamper both the knowledge acquisition process and the efficiency of reasoning. However, such concerns appear not to pose a significant hindrance to Influx$_1$ based on the following grounds.

- The former concern (regarding the number of numerical values required) is commonly discussed in the context of the practical application of (classical) probabilistic models. Such probabilistic models require either a global probability distribution, or joint probability distributions (e.g., Bayesian networks) where the number of numerical values required is *exponential* to the number of all variables involved (in the former case) or to the number of parents for the child (in each latter case). Such distributions are not required in Influx$_1$. Instead of requiring one to assume that any relation between two or more propositions is certain and deterministic, Influx$_1$ endows one with a capability to express his/her belief (by means of numerical values) when this is not the case. Therefore, the number of numerical values is only *linear* to the number of *uncertain/indeterministic* relations in the rule network (or the number of such parents for a child).

- With respect to the latter concern about the necessary precision of the numerical values, empirical experiments which studied the necessary precision of numerical values to measure uncertainty in quantitative reasoning (conducted with Bayesian networks [26, 40]) have shown insensitivity to the precision of those values. Thus approximations that distinguish different degrees of certainty, rather than precise reflections of reality, can be considered sufficient when using Influx$_1$ in many practical cases. For instance when a precise measure is not available, it might be possible for one to devise a scheme to approximate numerical measures of uncertainty based on qualitative measures such as

  $$\textit{impossible} < \textit{unlikely} < \ldots < \textit{maybe} < \textit{often} < \ldots < \textit{usually} < \textit{likely} < \textit{certainly}$$

  where $<$ denotes the relative degree of strength.

System dynamics Influx$_1$ has been developed to support dynamic applications. There is no pre-processing or initialisation phase, and changes to the system can be made at any time without incurring significant penalty. This is achieved by supporting changes to beliefs and knowledge in real-time, with forward and backward chaining used to propagate and construct beliefs and networks respectively in response. In this way, applications can be built that permit constant change and growth, as required for many real world situations.

Limitations of Influx$_1$ Despite the aforementioned qualities, Influx$_1$ has a number of limitations. Major limitations of Influx$_1$ are given below.

- Influx$_1$ provides a very flexible framework that offers diverse instruments to tailor a reasoning task to the available knowledge, specific requirements and specific situation at hand. However, this flexibility comes at a cost: the requirement that users have a sufficient understanding of the different methods and techniques provided by the tool with which to make correct judgments and decisions in constructing a rule network.

- Influx$_1$ does not yet possess capabilities that ease various modelling efforts. For instance, it is highly desirable for the tool to (i) automate the process of dependency analysis to assist the developers in choosing the appropriate combination/fusion operators (a task that may be demanding when the system is scaled to a large size with widespread dependencies), and (ii) automatically transfer belief between mutually exclusive propositions (a task that currently requires explicit modelling). The lack of these capabilities is felt most at scale, thus restricting scalability. The potential of automated dependency analysis further raises the prospect of supporting scenarios with dynamic dependency relationships that change at run-time, something not yet addressed in Influx$_1$.

- Influx$_1$ currently performs unidirectional reasoning (i.e., belief is propagated between propositions in one direction). In this respect, the reasoning power of an inferencing system can be enhanced by allowing for a bidirectional flow of information, thus achieving both types of reasoning (deductive and abductive) over the same knowledge structure. For instance, given $A \rightarrow B$, one can deductively derive B if A is verified, and conversely, can abductively infer A if B is observed. However, the same principle does not apply in the presence of uncertainty: given $A \xrightarrow{p} B$, one would encounter trouble predicting the belief for A solely based on the occurrence of B. In this regard, having both types of rule (e.g., $A \xrightarrow{p_1} B$ and $B \xrightarrow{p_2} A$) explicitly specified in the system would make bidirectional reasoning possible, but could cause cyclic inferencing, leading to beliefs being amplified and invalid conclusions being derived. The problem of cyclic inferencing can be eliminated by enabling the system to remove either of the rules when the other rule is activated [39], in which case the simplicity and elegance of the inferencing system inevitably diminishes. In practice, (traditional) expert systems dealing with uncertainty (such as the medical expert system MYCIN) tackle this problem by simply disallowing bidirectional flows of information [39].

The design of Influx$_2$ is motivated to rectify the aforementioned limitations and to augment the reasoning capability of Influx$_1$. More specifically, Influx$_2$ aspires to offer a seamless integration of deductive and abductive reasoning, as well as to enhance the coherency, usability and scalability of the reasoner, as well as to support dynamic dependency structure without significantly compromising the efficiency and simplicity of Influx$_1$.

References

1. Barnett, J. A. 1981. Computational methods for a mathematical theory of evidence. In Proceedings of the Seventh International Joint Conference on Artificial Intelligence (IJCAI), pp. 868–875.

2. Bauer, M. 1997. Approximation algorithms and decision making in the Dempster-Shafer theory of evidence — an empirical study. *International Journal of Approximate Reasoning*, Vol. 17, pp. 217–237.

3. Cano, J., Delgado, M. and Moral, S. 1993. An axiomatic framework for propagating uncertainty in directed acyclic networks. *International Journal of Approximate Reasoning*, pp. 253–280.

4. Daniel, M. 2010. Several notes on belief combination. Workshop on the Theory of Belief Functions, Brest, France.

5. Dempster, A. P. 1967. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, Vol. 38, No. 2, pp. 325–339, doi:10.1214/aoms/1177698950.

6. Denoeux, T. 2001. Inner and outer approximation of belief structures using a hierarchical clustering approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 9, No. 4, pp. 437-460.

7. Denoeux, T. 2006. The cautious rule of combination for belief functions and some extensions. In Proceedings of International Conference on Information Fusion, Florence, Italy.

8. Denœux, T. 2008. Conjunctive and disjunctive combination of belief functions induced by nondistinct bodies of evidence. *Artificial Intelligence*, vol. 172, pp. 234–264.

9. Dubois, D. and Prade, H. 1985. Combination and propagation of uncertainty with belief functions. In Proceedings of the Ninth International Joint Conference on Artificial Intelligence, pp. 111–113, Los Angles, California.

10. Dubois D. and Prade H. 1986, A Set Theoretical View of Belief Functions. *International Journal of General Systems*, Vol. 12, pp. 193-226.

11. Dubois D. and Prade H. 1988. Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, Vol. 4, pp. 244-264.

12. Dubois, D. and Prade, H. 1990. Resolution principles in possibilistic logic. *International Journal of Approximate Reasoning*, Vol. 4, No. 1, pp. 1–21.

13. Dubois, D., Lang, J. and Prade H. 1990. A possibilistic assumption-based truth maintenance system with uncertainty justifications, and its application to belief revision. In *Truth Maintenance Systems*, pp. 87–106, Springer Verlag.

14. Dubois, D. and Prade, H. 2009. Formal Representation of Uncertainty. Chapter 3 in *Decision-making Process: Concepts and Methods*. ISTE.

15. Falkenhainer, B. 1988. Towards a general-purpose belief maintenance system. *Uncertainty in Artificial Intelligence 2*, pp. 125–131.

16. Fan, X. and Zuo, J. M. 2006. Fault diagnosis of machines based on D-S evidence theory. Part 1: D-S evidence theory and its improvement. *Pattern Recognition Letters 27*, pp. 366–376. Elsevier.

17. Garvey, T. D., Lowrance, J. D. and Fischler, M. A. 1981. An inference technique for integrating knowledge from disparate sources. In Proceedings of the Seventh International Joint Conference in Artificial Intelligence, pp. 319–325.

18. Ginsberg, M. L. 1984. Non-monotonic reasoning using Dempsters rule. In Proceedings of the Fourth National Conference on Artificial Intelligence (AAAI-84), TX, US, pp. 126–129.

19. Gordon, J. and Shortliffe, E. H. 1984. The Dempster-Shafer theory of evidence: Rule-based expert systems. The MYCIN Experience of the Stanford Heuristic Programming Project, pp. 272–292. Addison Wesley, Reading, MA.

20. Gordon, J. and Shortliffe, E. H. 1985. A method for managing evidential reasoning in hierarchical hypothesis spaces. *Artificial Intelligence*, Vol. 26, pp. 323–358.

21. Haenni, R. 2004. Ordered valuation algebras: a generic framework for approximate inference. *International Journal of Approximate Reasoning*, Vol. 37,inference, Int. Journal of Approximate Reasoning, vol. 37, pp. 1–41.

22. Haenni, R., Kohlas, J. and Lehmann, N. 2000. Probabilistic argumentation systems. *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol. 5: Algorithms for Uncertainty and Defeasible Reasoning, pp. 221-287

23. Haenni, R. and Lehmann, N. 2002. Resource bounded and anytime approximation of belief function computations. *International Journal of Approximate Reasoning*, Vol. 31, pp. 103-154.

24. Haenni, R. and Lehmann, N. 2003. Probabilistic argumentation systems: a new perspective on the Dempster-Shafer theory. *International Journal of Intelligence System*, Vol. 18, pp. 93–106.

25. Haenni, R. 2005. Shedding new light on Zadeh's criticism of Dempster's rule of combination. In Proceedings of the Eighth International Conference on Information Fusion, Philadelphia, PA, USA.

26. Henrion, M., Pradhan, M., del Favero, B., Huang, K., Provan, G. and O'Rorke, P. 1996. Why is diagnosis using belief networks insensitive to imprecision in probabilities? In Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence, San Francisco, CA, pp. 307–314.

27. Jøsang, A. 1997. Artificial Reasoning with subjective logic. In Proceedings of the 2nd Australian Workshop on Commonsense Reasoning, Perth. Australian Computer Society.

28. Jsang, A. 2008. Cumulative and Averaging Unfusion of Beliefs. In The Proceedings of the International Conference on Information Processing and Management of Uncertainty, Malaga, Spain.

29. Kennes, R. and Smets, Ph. 1991. Computational aspects of the Möbiũs transform. In Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence (UAI1990), Amsterdam, pp. 401-416.

30. Laskey, K. B. and Lehner, P. E. 1990. Assumptions, beliefs and probabilities. *Artificial Intelligence*, Vol. 32, pp. 65–77.

31. Lehmann, N. 2001. Argumentation Systems and Belief Functions. PhD Thesis, Freiburg University.

32. Lowrance, H., Garvey, T. and Strat, T. 1986. A framework for evidential-reasoning systems. In Proceedings of the 5th National Conference on Artificial Intelligence (AAAI'86), Vol. 2, Morgan Kaufmann, pp. 896-903.

33. Łukasiewicz, J. 1967. *On 3-valued logic*. Oxford University Press, Oxford, pp. 15–18.

34. McCarthy, J. 1980. Circumscription: A form of non-monotonic reasoning. *Artificial Intelligence*, Vol. 13, No. 1–2, pp. 27–39.

35. Murphy C. K. 2000. Combining belief functions when evidence conflicts. *Decision Support Systems*, Vol. 29, pp. 1-9.

36. Nilsson, N. 1986. Probabilistic Logic. *Artificial Intelligence*, Vol. 28, No. 1, pp. 71–87.

37. Oxenham, M. 2008. The effect of finite set representations on the evaluation of Dempsters rule of combination. In Proceedings of the International Conference on Information Fusion, Germany, pp. 1-8.

38. Parson, S. and Hunter, A. 1998. A Review of Uncertainty Handling Formalisms. *Applications of Uncertainty Formalisms*, *LNCS*, Vol. 1455, pp. 266–302. Springer-Verlag.

39. Pearl, J. 1988. *Probabilistic Reasoning Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers.

40. Pradhan, M., Henrion, M., Provan, G., de Favero, B. and Huang, K. 1996. The sensitivity of belief networks to imprecise probabilities: An experimental investigation. *Artificial Intelligence*, Vol. 85, pp. 363–397.

41. Provan, G. M. 1988. Solving diagnostic problems using extended assumption-based truth maintenance system: foundations. Technical Report 88-10, Department of Computer Science, University of British Columbia.

42. Quost, B., Denoeux, T. and Mylene M. 2006. One-against-all classifier combination in the framework of belief functions. In Proceedings of the Eighth Conference on Information Fusion Conference, pp. 356–363.

43. Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence*, Vol. 13, No. 1–2, pp. 81–132.

44. Sentz, K. and Ferson, S. 2002. Combination of Evidence in Dempster-Shafer Theory. Sandia report. Sandia National Laboratories.

45. Shackle, G. L. S. 1961. *Decision, order and time in human affairs.* Cambridge University Press, Cambridge, UK.

46. Shafer, G. 1976. *A Mathematical Theory of Evidence*, ISBN 0-608-02508-9, Princeton University Press.

47. Shafer, G. and Logan, R. 1987. Implementing Dempster's rule for hierarchical evidence. *Artificial Intelligence*, Vol. 33, pp. 271–298.

48. Smarandache, F. and Dezert, J. 2005. Information Fusion Based on New Proportional Conflict Redistribution Rules. In Proceedings of Fusion 2005 International Conference on Information Fusion, Philadelphia, US.

49. Smets, P. 1993. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of Approximate Reasoning*, Vol. 9, No. 1, pp. 1–35. Elsevier.

50. Smets, P. 1994. What is Dempster-Shafer's Model?. *Advances in the Dempster-Shafer Theory of Evidence*, pp. 5–34. Wiley.

51. Smets, P. and Kennes, R. 1994. The transferable belief model. *Artificial Intelligence*, Vol. 66, pp. 191–234.

52. Smithson, M. 1989. *Ignorance and Uncertainty: Emerging Paradigms.* Springer Verlag, New York.

53. Shenoy, P. P. and Shafer, G. 1986. Propagating belief functions with local computations. *IEEE Expert*, Vol. 1, No. 3, pp. 43–52.

54. Shenoy, P. P. and Shafer, G. 1990. Axioms for probability and belief-function propagation. *Uncertainty in Artificial Intelligence*, Vol. 4, pp. 169–198.

55. Srinivas, S. 1994. *A Probabilistic ATMS.* Knowledge systems laboratory, Standford University.

56. Suter, G., Barnthouse, L. and O'Neil, R. 1987. Treatment of risk in environmental impact assessment. *Environmental Management*, Vol. 11, No. 3, pp. 295–303. Springer-Verlag.

57. Tessem, B. 1993. Approximations for efficient computation in the theory of evidence. *Artificial Intelligence*, Vol. 61, No. 2, pp. 315-329.

58. van Dam, K. 1998. Using uncertainty techniques in radio communication systems. *Applications of Uncertainty Formalisms*, pp. 76–87, Springer Verlag.

59. Voorbraak, F. 1989. A computationally efficient approximation of DempsterShafer theory. *International Journal of ManMachine Studies*, Vol. 30, No. 5, pp. 525-536.

60. Walker, W., Harremoés, P., Rotmans, J., van der Sluijs, J., van Asselt, M., Janssen, P. and Krayer von Krauss, M. 2003. Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, Vol. 4, No. 1, pp. 5–18. Swets & Zeitlinger.

61. Wilson, N. 1991. A Monte-Carlo algorithm for Dempster-Shafer belief. In Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence, US, pp. 414-417.

62. Wilson, N. and Moral, S. 1996. Fast Markov chain algorithms for calculating Dempster-Shafer belief. In Proceedings of the Twelfth European Conference on Artificial Intelligence, Hungary, pp. 672-678.

63. Wynne, B. 1992. Uncertainty and environmental learning: Reconceiving science and policy in the preventive paradigm. *Global Environmental Change*, Vol. 2, pp. 111–127. Butterworth-Heinemann Ltd.

64. Xu, H. and Smets, P. 1996. Reasoning in Evidential Networks with Conditional Belief Functions. *International Journal of Approximate Reasoning*, Vol. 14, pp. 155–185. Elsevier.

65. Yager R. R. 1983. Hedging in the combination of evidence. *Journal of Information Optimization Sciences*, Vol. 4, No. 1, pp. 73-81.

66. Yaghlane, A. B., Denoeux, T. and Mellouli, K. 2001. Coarsening approximations of belief functions. In Proceedings of Symbolic and Quantitative Approaches to Reasoning with Uncertainty, pp. 362–373.

67. Yaghlane, B. B. and Mellouli, K. 2008. Inference in directed evidential networks based on the transferable belief model. *International Journal of Approximate Reasoning*, Vol. 48, pp. 399–418. Elsevier.

68. Zadeh, L. A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, Vol. 1, pp. 1–28.

69. Zadeh, L. A. 1979. On the validity of Dempsters rule of combination of evidence. Technical Report 79/24, University of California, Berkely, US.

70. Zadeh, L. A. 1984. A mathematical theory of evidence (book review). *AI Magazine*, Vol. 55, pp. 81–83.

71. Zadeh, L. A. 1986. A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination. *AI Magazine*, Vol. 7, pp. 85–90.

DEFENCE SCIENCE AND TECHNOLOGY GROUP DOCUMENT CONTROL DATA	1. DLM/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
Influx$_1$: A Tool and Framework for Reasoning under Uncertainty	Document (U) Title (U) Abstract (U)

4. AUTHORS	5. CORPORATE AUTHOR
Van Nguyen and Michael Docking	Defence Science and Technology Group PO Box 1500 Edinburgh, South Australia 5111, Australia

6a. DST GROUP NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DST Group–TR–3142	016-372	Technical Report	September, 2015

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. No. OF PAGES	12. No. OF REFS
			72	71

13. DST Group Publications Repository	14. RELEASE AUTHORITY
`http://dspace.dsto.defence.gov.au/dspace/`	Chief, Cyber and Electronic Warfare Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for Public Release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SOUTH AUSTRALIA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS

No Limitations

18. DST GROUP RESEARCH LIBRARY THESAURUS

reasoning under uncertainty, information fusion, Dempster-Shafer theory, theory of belief functions, quantitative reasoning

19. ABSTRACT

Influx is a reasoning component of REALISE, a generic platform for intelligent systems currently under investigation as part of an ongoing research program conducted by the CAO Branch. Influx is envisioned to have two major uses: facilitating and mediating reasoning processes for intelligent missions, and serving as a tool and framework for reasoning under uncertainty. This document reports on some initial research and development efforts pertaining to the reasoning aspect of Influx in the latter scenario. Due in part to its generality, the Dempster-Shafer (D-S) theory is chosen as a theoretical basis for representing imperfect knowledge and for reasoning with such knowledge. Since Influx is intended to deal with real-time and real-life applications, it is of primary importance for the reasoning tool to practically achieve adequate performance, flexibility, scalability and system dynamics. To this end, Influx aspires to reach such objectives while attaining a plausible reasoning mechanism formulated from D-S methods and techniques. The initial version, Influx$_1$, is a simple, but highly efficient and flexible, nonmonotonic rule-based system enhanced with D-S based belief representation, fusion and inference. Influx$_1$ has been applied towards tasks that include situational awareness and network traffic analysis. This document provides a high-level description of Influx$_1$ from the reasoning perspective. Research and development pertaining to the implementation of the reasoning tool and specific applications are not included in this document.