An Aleatoric Description Logic for Probabilistic Reasoning

Tim French[1[0000-0002-0748-8040]] and Thomas Smoker[1]

The University of Western Australia, tim.french@uwa.edu.au
thomas.smoker@research.uwa.edu.au

Abstract. Description logics are a powerful tool for describing ontological knowledge bases. That is, they give a factual account of the world in terms of individuals, concepts and relations. In the presence of uncertainty, such factual accounts are not feasible, and a subjective or epistemic approach is required. Aleatoric description logic models uncertainty in the world as aleatoric events, by the roll of the dice, where an agent has subjective beliefs about the bias of these dice. This provides a subjective Bayesian description logic, where propositions and relations are assigned probabilities according to what a rational agent would bet, given a configuration of possible individuals and dice. Aleatoric description logic is shown to generalise the description logic ALC, and can be seen to describe a probability space of interpretations of a restriction of ALC where all roles are functions. Several computational problems are considered and aleatoric description logic is shown to be able to model learning, via Bayesian conditioning.

Keywords: Probabilistic Reasoning · Belief Representation · Learning Agents

1 Introduction

Description logics [1] give a formal foundation for ontological reasoning: reasoning about factual aspects of the world. However, many reasoning tasks are performed in the presence of incomplete or uncertain information, so a reasoner must apply some kind of belief model to approximate the true state of the world. This work investigates the application of description logics to describing uncertain and incomplete concepts, following the recent development of aleatoric modal logic [7]. The term aleatoric has its roots in the Latin *aleator*, meaning dice player, and it is this origin that motivates this work. Concepts are not simply described as matters of fact, but can be more generally described as *reasonable bets*. While a person may definitely have a virus or not, a virus test kit with 95% accuracy is effectively a role of a dice. That is, on receiving a positive test a rational person would accept a 20-1 bet that they're not infected. Therefore concepts may be modelled using probabilities corresponding to a rational bet that the concept holds true, along the lines of the Dutch book argument of Ramsey [21] and de Finetti [5]. The fundamental assumption of this work is that an agent models the world *aleatorically*, where events correspond to the roll of dice, and the bias of the dice is treated *epistemically*. That is, the agent has prior assumptions about the bias of the dice, and may refine these assumptions through observing the world.

Aleatoric description logic aims to model reasoning in uncertain and subjective knowledge settings [10, 12]. However, aleatoric description logic takes an approach where the probabilistic and logical aspects of the knowledge base are completely unified, rather than several other approaches where these are independent facets of the knowledge base [4, 23, 17]. Therefore all concept and roles are represented by "dice rolls" corresponding to an agent's beliefs on the likely configuration of the world.

An advantage of this "probability first" approach is that aleatoric description logic is naturally able to model learning via Bayesian conditioning over complex observations (i.e. logical formula). Aleatoric modal logic is introduced in [7], where the semantics are presented along with a proof theoretic calculus. This paper extends that syntax and semantics to an aleatoric description logic, following the correspondence between description logics and modal logics [1].

2 Aleatoric Description Logic

This section presents the core syntax and semantics for Aleatoric Description Logic (ADL). This is a generalisation of standard description logics, such as $\mathcal{ALC}$, in the same sense that complex arithmetic is a generalisation of real-valued arithmetic: the true-false/0-1 values of description logics are extended to the closed interval $[0, 1]$.

The syntax of ADL varies from that of $\mathcal{ALC}$ in a number of ways: there is a ternary operator, *if-then-else*, in place of the typical Boolean operators, and a *marginalisation* operator in place of the normal role quantifiers. These operators add expressivity, but also better capture the aleatoric intuitions of the logic.

The syntax for ADL is specified with respect to a set of *atomic concepts*, X and a set of *roles*, R:

$$\alpha \ ::= \ \top \mid \bot \mid A \mid \ (\alpha?\alpha{:}\alpha) \ \mid [\rho]\,(\alpha\,|\,\alpha)$$

where $A \in \mathrm{X}$ is an atomic concept, and $\rho \in \mathrm{R}$ is a role. Let the set of ADL formulas generated by this syntax be $\mathcal{L}_{\mathrm{ADL}}$. This syntax uses non-standard operators and the following terminology is used: $\top$ is *always*; $\bot$ is *never*; A is some named concept that may hold for an individual; $(\alpha?\beta{:}\gamma)$ is *if* α *then* β, *else* γ; and $[\rho]\,(\alpha\,|\,\beta)$ is ρ *is* α *given* β.

We also identify a special role $\mathtt{id} \in \mathrm{R}$ referred to as *identity*, which essentially refers to different possibilities for the one individual, and write $(\alpha\,|\,\beta)$ in place of $[\mathtt{id}]\,(\alpha\,|\,\beta)$.

In these semantics every thing is interpreted as a probability dependent only on the individual: $\top$ always has probability 1.0 and $\bot$ always has probability

0.0; an atomic concept A has some probability that is dependent only on the current individual; $(\alpha?\beta{:}\gamma)$ has the probability of β given α or γ given not α; and $[\rho]\,(\alpha\,|\,\beta)$ is the probability of α given β over the set of individuals in the probability distribution corresponding to ρ.

2.1 Probabilistic Semantics

A formula of aleatoric description logic is interpreted with respect an *aleatoric belief model*, that is based on the probability model of [7] and the probability structures defined in [10]. Given a countable set S, we use the notation $\mathrm{PD}(S)$ to notate the set of *probability distributions* over S, where $\mu \in \mathrm{PD}(S)$ implies: $\mu : S \longrightarrow [0,1]$; and $\sum_{s\in S}\mu(s) = 1$.

Definition 1. *Given a set of atomic concepts* X*, and a set of roles* R*, an* aleatoric belief model *is specified by the tuple* $\mathcal{B} = (I, r, \ell)$*, where:*

- I *is a set of possible individuals.*
- $r : \mathrm{R} \times I \longrightarrow \mathrm{PD}(I)$ *assigns for each role* $\rho \in \mathrm{R}$ *and each individual* $i \in I$*, a probability distribution* $r(\rho, i)$ *over* I*. We will typically write* $\rho(i,j)$ *in place of* $r(\rho,i)(j)$*.*
- *For the role* `id`*, we include the additional constraint: for all* $i,j,k \in I$*,* $\mathtt{id}(i,j) > 0$ *implies* $\mathtt{id}(j,k) = \mathtt{id}(i,k)$*.*
- $\ell : I \times \mathrm{X} \longrightarrow [0,1]$ *gives the likelihood,* $\ell(i, C)$ *of an individual* i *satisfying an atomic concept* C*. We will write* $C(i)$ *in place of* $\ell(C,i)$*.*

Given some $i \in I$*, we let* $\mathcal{B}_i$ *be referred to as a* pointed aleatoric belief model.

Definition 2. *Given an aleatoric belief model* $\mathcal{B} = (I, r, \ell)$*, some* $i \in I$*, and some* $\alpha \in \mathrm{ADL}$ *we specify the probability* $\mathcal{B}$ *assigns* i *satisfying* α*,* $\mathcal{B}_i(\alpha)$*, recursively. We use the abbreviation* $E_i^\rho\alpha = \sum_{j\in I}\rho(i,j)\mathcal{B}_j(\alpha)$*, where* $\rho \in \mathrm{R}$*. Then:*

$$\begin{array}{ll}
\mathcal{B}_i(\bot) = 0 & \mathcal{B}_i((\alpha?\beta{:}\gamma)) = \mathcal{B}_i(\alpha).\mathcal{B}_i(\beta) + (1 - \mathcal{B}_i(\alpha)).\mathcal{B}_i(\gamma) \\
\mathcal{B}_i(\top) = 1 & \mathcal{B}_i([\rho]\,(\alpha\,|\,\beta)) = \frac{\sum_{j\in I}\rho(i,j)\mathcal{B}_j(\alpha)\mathcal{B}_j(\beta)}{E_i^\rho\beta} \ \textit{if}\ E_i^\rho\beta > 0 \\
\mathcal{B}_i(C) = C(i) & \mathcal{B}_i([\rho]\,(\alpha\,|\,\beta)) = 1,\ \textit{if}\ E_i^\rho\beta = 0
\end{array}$$

In these semantics each proposition can be seen as an event (or a series of conditional events), and the interpretation describes the probability of that event being observed.

For example, the following proposition describes the concept of someone having been exposed to a virus, given they were in contact with someone who had a fever

$$\mathit{exp} = (\mathit{virus}?\top{:}[\mathit{contact}]\,(\mathit{infectious}|\mathit{fever})) \tag{1}$$

"The person was either already (asymptomatically) infected, or some person selected from the population of contacts who have a fever, was infectious".

To evaluate this the concept *virus* is sampled, to see if they already have the virus (there may be a 1% chance). In the cases where they didn't already have the virus, random individuals are sampled from the population of contacts, and the concept *fever* is sampled from the individuals until a febrile contact is selected. The probability associated to the proposition *exp* is the probability of this process selecting an individual where *infectious* is sampled.

An important property of these semantics is the weak independence assumption: *All formulas of* ADL *are contingent only on the individual at which they are evaluated*. This means that two formulas evaluated at the same individual may be viewed as independent probabilistic events.

That is, the probability of a coin landing heads twice in a row is independent of the probability of the same coin landing heads once. Both events are contingent on the bias of the coin, so in universes where the coin is more likely to land heads, both events are more likely, but the event are conditionally independent given the universe (or the possible individual). This simplifies the representation of complex dependencies, as the joint probability of all events can be constrained to be a probability distribution of individuals, where all event are conditionally independent given an individual.

2.2 Aleatoric Knowledge Bases

An aleatoric knowledge base is defined over the same signature of atomic concepts X and roles R, including `id`. Additionally there is a set of *named individuals*, N, which may be thought of as special concepts for grounding assertions and framing queries. In line with the epistemic nature of these knowledge bases each named individual can be any one of a number of *possible individuals*, and the distribution of these possible individuals is represented by the role `id`.

As with $\mathcal{ALC}$ we have terminological axioms and assertional axioms. The *aleatoric terminological axioms* or *T-Books* describe rules that are universally true for all individuals, and thus provide a non-probabilistic intensional definition of the concepts and roles in the logic. *Aleatoric assertional axioms* or *A-Books* describe subjective extensional information by listing the probabilities with which individuals satisfy given concepts and roles. It is not the case that *T-Books* describe *concept inclusion* nor *subsumption* as *TBoxes* do in $\mathcal{ALC}$, as these concepts do not have a strong intuitive foundation in an aleatoric setting. Instead *T-Books* provide a means to *constrain strength of belief*.

Definition 3. *The* aleatoric terminological axioms *have the form:*

$$\alpha \preceq \beta \ (\alpha \textit{ is no more likely than } \beta) \quad \textit{and} \quad \alpha \approx \beta \ (\alpha \textit{ is exactly as likely as } \beta).$$

A T-Book *is a set of aleatoric terminological axioms.*

These axioms place universal constraints on the likelihoods of aleatoric formulas being true. For example we might include an axiom *first* $\preceq$ *place*, meaning coming first in a race is no more likely than placing (coming first, second or third). Alternatively, we could define placing precisely as coming first, second or third, via the axiom *place* $\approx$ *first* $\sqcup$ *second* $\sqcup$ *third*, and then *first* $\preceq$ *place* is implicitly true.

Definition 4. *The* aleatoric assertional axiom *(or simply assertions) have the form:*

- $a :^p \alpha$*, where* $a \in \mathrm{N}$*,* $p \in [0,1]$ *and* $\alpha \in$ ADL *asserts a belief that individual* a *satisfies concept* α*, with probability* p*.*
- $(a,b) :^p \rho$*, where* $a, b \in \mathrm{N}$*,* $p \in [0,1]$ *and* $\rho \in \mathrm{R}$ *asserts that individual* b *satisfies the role* ρ *for* a *with probability* p*.*

An A-Book $\mathcal{A}$ *is a set of aleatoric assertional axioms, and* $\mathcal{A}$ *is a* well-formed A-Book *if for every* $a \in N$*, for every* $\rho \in R$*,* $\sum_{b\in\mathrm{N}}\{p \mid (a,b) :^p \rho\} \leq 1$*.*

While T-Books give intensional definitions of concepts, A-Books give an extensional definition by describing the concepts and roles in terms of the individuals. As such an A-Book captures an agent's current state of belief, just as a bookmaker's book describes the current odds for a race. An A-Book is well-formed if the roles described can be extended to a probability distribution (i.e. the probabilities sum to less than 1).

While an A-Book is existentially quantified, T-Books are universally quantified and consequently a very powerful formalism.

Definition 5. *An* aleatoric knowledge base $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ *is a pair consisting of a set of assertional axioms* $\mathcal{A}$ *and a set of terminological axioms* $\mathcal{T}$*.*

An aleatoric knowledge base describes a belief, or subjective position of an agent, that can correspond to a number of different interpretations. The semantics for these interpretations are given below.

An interpretation *satisfies* the aleatoric knowledge base $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ iff it satisfies all the axioms in $\mathcal{A}$ and $\mathcal{T}$.

Definition 6. *Given an aleatoric knowledge base* $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ *over the signature* $(\mathrm{X}, \mathrm{R}, \mathrm{N})$*, and an aleatoric belief model* $\mathcal{B} = (I, R, \ell)$ *over the signature* $(\mathrm{X}\cup\mathrm{N}, \mathrm{R})$ satisfies $\mathcal{K}$ *iff:*

- *For every* $a \in \mathrm{N}$*, for every* $i \in I$*,* $a(i) \in \{0,1\}$ *and for all* $i, j \in I$*,* $a(i) = 1$ *and* $\mathtt{id}(i,j) > 0$ *implies* $a(j) = 1$*. That is, the names are absolute concepts, and two possibilities for a single individual will share a name.*
- *For axioms* $\alpha \preceq \beta \in \mathcal{T}$*, for all* $i \in I$*,* $\mathcal{B}_i(\alpha) \leq \mathcal{B}_i(\beta)$*.*
- *For axioms* $\alpha \approx \beta \in \mathcal{T}$*, for all* $i \in I$*,* $\mathcal{B}_i(\alpha) = \mathcal{B}_i(\beta)$*.*
- *For axioms* $a :^p \alpha \in \mathcal{A}$*, for all* $i \in I$ *with* $a(i) = 1$*,* $\mathcal{B}_i(E\alpha) = p$*.*
- *For axioms* $(a,b) :^p \rho \in \mathcal{A}$ *for all* $i \in I$ *with* $a(i) = 1$*,* $\sum_{j\in I} \rho(i,j) \cdot b(j) = p$*.*

We say that a knowledge base $\mathcal{K}$ *is* consistent *if it is supported by at least one aleatoric belief model.*

Table 1 gives a set of abbreviations familiar in the context of description logics.

The abbreviation in the right column of the table corresponds to a process of repeated sampling: Where $n, m \in \omega$, $\alpha^{\frac{n}{m}}$ corresponds to the likelihood of α being sampled at least n times out of m. (A similar abbreviation can be defined

Table 1: Some abbreviations of operators in ADL.

term	**formula**	**interpretation**
$\alpha \sqcap \beta$	$(\alpha?\beta{:}\bot)$	$\mathcal{B}_i(\alpha).\mathcal{B}_i(\beta)$
$\alpha \sqcup \beta$	$(\alpha?\top{:}\beta)$	$\mathcal{B}_i(\alpha) + \mathcal{B}_i(\beta) - \mathcal{B}_i(\alpha).\mathcal{B}_i(\beta)$
$\neg\alpha$	$(\alpha?\bot{:}\top)$	$1 - \mathcal{B}_i(\alpha)$
$\alpha \Rightarrow \beta$	$(\alpha?\beta{:}\top)$	$1 - \mathcal{B}_i(\alpha) + \mathcal{B}_i(\alpha).\mathcal{B}_i(\beta)$
$E_\rho\alpha$	$[\rho]\,(\alpha\,\vert\,\top)$	$\sum_{j\in I} \rho(i,j).\mathcal{B}_j(\alpha)$
$\exists\rho.\alpha$	$\neg[\rho]\,(\bot\,\vert\,\alpha)$	1 if $E_\rho\alpha \neq 0$, 0 otherwise.

$$\alpha^{\frac{n}{m}} = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } m < n \\ \left(\alpha?\alpha^{\frac{n-1}{m-1}}{:}\alpha^{\frac{n}{m}}\right) & \text{if } n < m \end{cases}$$

for α coming up $\top$ *exactly* n times out of m.) Note that this does not describe a probability or frequency, but an event. So $\alpha^{\frac{4}{5}}$ does not mean α is sampled at least 80% of the time. Instead it describes the event of α being sampled 4 times out of 5, which would be quite likely (0.88) if α had probability 0.8, and unlikely (0.19) if α had probability 0.5. This formalism can encode degrees of belief in an elegant way. If an agent were to perform an action only if they believed α very strongly, one might set $\alpha^{\frac{9}{10}}$ as a precondition for the action, and if an agent were informed of a proposition β by another agent who is considered unreliable, they may update their belief base with the proposition $\beta^{\frac{2}{3}}$.

These operators may not appear logical: $\sqcap$ is not idempotent, and appears similar to the product t-norm of fuzzy logic [25]. However, Section 3 shows that these new operators are inherently probabilistic and represent the process of reasoning over a probability space of description logic models. Furthermore, restricting the concept probabilities to be 0 or 1, it can be seen that the semantic interpretation of $\sqcap$, $\neg$ and $\exists\rho$ agrees with the standard description logic semantics, so classical description logic can be seen as a special case of aleatoric description logic.

2.3 Example

For example, suppose we have three agents: `Hector`, `Igor` and `Julia`. They each may have a virus (V), or not, and they also may have a fever (F), whether they have the virus or not. For each possible individual, there is a probability of them having a fever, which is naturally higher for possible individuals with the virus. Each agent will occasionally come into contact with another agent, and the identity of this agent is described by the probability distribution `contact`. Finally, for each possible individual there is the probability of them being the actual agent (`id`).

An aleatoric knowledge base could model that Hector is very likely not to have the virus; Julia is likely to have the virus, Julia is very likely to have a fever and it is likely that Hector came into contact with Julia. Furthermore, a terminological axiom can specify the belief that a new exposure to the virus (`exp`) occurs if an agent did not already have the virus, but came into contact with some febrile person who did have the virus. We can calculate the probability of an agent being newly exposed to the virus:

$$E(\neg V \sqcap [\mathtt{c}]\,(V\,|\,F)) \preceq \mathtt{exp}$$

Thus the aleatoric knowledge base $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ is:

$$\mathcal{A} = \left\{\mathtt{Hector} :^{0.1} V,\ \mathtt{Julia} :^{0.7} V,\ \mathtt{Julia} :^{0.69} F,\ (\mathtt{Hector}, \mathtt{Julia}) :^{0.3} \mathtt{c}\right\}$$
$$\mathcal{T} = \{E(\neg V \sqcap [\mathtt{c}]\,(V\,|\,F)) \preceq \mathtt{exp}\}$$

To determine if the knowledge base necessitates that there is a greater than 25% chance of Hector being newly exposed to the virus, the axiom $\mathtt{Hector} :^{0.25} \mathtt{exp}$ can be inserted into the knowledge base, and consistency checking can be applied. This process is described in the following section.

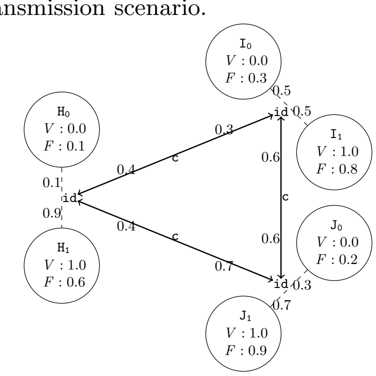
An interpretation that satisfies the knowledge base is presented below. For each agent, we suppose that there are two possible individuals (**PI**), one with the virus (e.g. `Hector1`) and one without (e.g. `Hector0`). Note that the weighted probabilities of these agents satisfy the constraints of the A-Book, $\mathcal{A}$.

The probabilities for this scenario are given in Table 2, and a graphical representation is given in Figure 1.

Table 2: Initial probabilities for agent, contacts, virus and symptoms

PI	id	V	F	$\mathtt{H}_0$	$\mathtt{H}_1$	$\mathtt{I}_0$	$\mathtt{I}_1$	$\mathtt{J}_0$	$\mathtt{J}_1$
$\mathtt{Hector}_0$	0.9	0.0	0.1	0.0	0.0	0.15	0.15	0.21	0.49
$\mathtt{Hector}_1$	0.1	1.0	0.6	0.0	0.0	0.15	0.15	0.21	0.49
$\mathtt{Igor}_0$	0.5	0.0	0.3	0.04	0.36	0.0	0.0	0.18	0.42
$\mathtt{Igor}_1$	0.5	1.0	0.8	0.04	0.36	0.0	0.0	0.18	0.42
$\mathtt{Julia}_0$	0.3	0.0	0.2	0.04	0.36	0.3	0.3	0.0	0.0
$\mathtt{Julia}_1$	0.7	1.0	0.9	0.04	0.36	0.3	0.3	0.0	0.0

Fig. 1: A graphical example of the virus transmission scenario.



Interpreting this for `Hector`, we see the probability `Hector` was newly exposed to the virus is approximately 0.7. The working for this is shown in Table 3.

2.4 Reasoning with Aleatoric Description Logic

This section will consider computational properties of aleatoric description logic. The particular questions considered are:

Model Checking: Given an aleatoric belief model $\mathcal{B}_i$ and some formula α, what is the value of $\mathcal{B}_i(\alpha)$?

Satisfiability: Given an aleatoric knowledge base, $\mathcal{K} = (\mathcal{A}, \mathcal{T})$, is it consistent?

Table 3: A calculation of the chance of Hector being newly exposed with the virus, after a chance encounter with a person with a fever.

$\mathtt{I}_0$	$F_{\mathtt{I}_0} = 0.3, \quad (V \sqcap F)_{\mathtt{I}_0} = 0.0, \quad \mathtt{c}(\mathtt{H}_0, \mathtt{I}_0) = \mathtt{c}(\mathtt{H}_1, \mathtt{I}_0) = .15$
$\mathtt{I}_1$	$F_{\mathtt{I}_1} = 0.8, \quad (V \sqcap F)_{\mathtt{I}_0} = 0.8, \quad \mathtt{c}(\mathtt{H}_0, \mathtt{I}_0) = \mathtt{c}(\mathtt{H}_1, \mathtt{I}_0) = .15$
$\mathtt{J}_0$	$F_{\mathtt{J}_0} = 0.2, \quad (V \sqcap F)_{\mathtt{J}_0} = 0.0, \quad \mathtt{c}(\mathtt{H}_0, \mathtt{J}_0) = 0.21, \quad \mathtt{c}(\mathtt{H}_1, \mathtt{J}_0) = .49$
$\mathtt{J}_1$	$F_{\mathtt{J}_1} = 0.9, \quad (V \sqcap F)_{\mathtt{J}_0} = 0.9, \quad \mathtt{c}(\mathtt{H}_0, \mathtt{J}_0) = 0.21, \quad \mathtt{c}(\mathtt{H}_1, \mathtt{J}_0) = .49$
$\mathtt{H}_0$	$V_0 = 0.0, \quad \mathtt{id}_0 = 0.9, \quad [\mathtt{c}]\,(V \,\vert\, F) = \frac{\sum_{x=\mathtt{I}_0}^{\mathtt{J}_1} \mathtt{c}(\mathtt{H}_0,x)\cdot(V \sqcap F)_x}{\sum_{x=\mathtt{I}_0}^{\mathtt{J}_1} \mathtt{c}(\mathtt{H}_0,x).F_x} = 0.78$
$\mathtt{H}_1$	$V_1 = 1.0, \quad \mathtt{id}_1 = 0.1, \quad [\mathtt{c}]\,(V \,\vert\, F) = \frac{\sum_{x=\mathtt{I}_0}^{\mathtt{J}_1} \mathtt{c}(\mathtt{H}_1,x)\cdot(V \sqcap F)_x}{\sum_{x=\mathtt{I}_0}^{\mathtt{J}_1} \mathtt{c}(\mathtt{H}_1,x)\cdot F_x} = 0.78$
$\mathtt{H}$	$E(\neg V \sqcap [\mathtt{c}]\,(V \,\vert\, F)) = \sum_{x=0}^{1}(1 - V_x) \cdot \mathtt{id}_\mathtt{x}.[\mathtt{c}]\,(V \,\vert\, F) = 0.7$

The main question of interest is whether there is any interpretation that could possibly correspond to a given aleatoric knowledge base. However, by assigning flat priors to all unknown (or *ambivalent* concepts) one can define an interpretation and get a partial answer via model-checking.

Theorem 1. *Given a pointed belief model $\mathcal{B}_i$ consisting of n possible individuals, and a formula α consisting of m symbols, the value $\mathcal{B}_i(\alpha)$ can be computed in time $O(n^2 m)$.*

See [8] (Lemma 4.7) for proof.

To be able to perform inference based on an aleatoric knowledge base, we must first determine if it is consistent (i.e. agrees with at least one aleatoric belief model). A partial solution is given here for *acyclic aleatoric knowledge bases.*

Definition 7. *A concept C is an* atom *if $C \in \mathrm{X} \cup \{\top, \bot\}$ (i.e. C is an atomic concept, always, or never). A terminological axiom is* simple *if it has the one of the forms*

- $C \approx (D?E{:}F)$ *where C, D, E and F are all atoms.*
- $C \approx [\rho]\,(D \,|\, E)$ *where C, D and E are all atoms.*

A simple T-Book, *$\mathcal{T}$, is a T-Book consisting only of simple terminological axioms. A simple T-Book, $\mathcal{T}$, is* acyclic *if there is no sequence of concepts $C_0, \ldots, C_n$ where:*

- *for all $i = 1, \ldots, n$, either:*
  - *there is some $C \approx (D?E{:}F) \in \mathcal{T}$, where $C_i, C_{i-1} \in \{C, D, E, F\} \cap \mathrm{X}$;*
  - *there is some $C \approx [\rho]\,(D \,|\, E) \in \mathcal{T}$, where $C_i, C_{i-1} \in \{C, D, E\} \cap \mathrm{X}$;*
- *there is some i where $C \approx [\rho]\,(D \,|\, E) \in \mathcal{T}$ and $C_i, C_{i-1} \in \{C, D, E\} \cap \mathrm{X}$;*
- *$C_0 = C_n$.*

An A-Book, *$\mathcal{A}$ is* simple *if for all aleatoric assertional axioms $\sigma \in \mathcal{A}$ of the form $a :^p \alpha$, it is the case that α is an atomic concept. If $\mathcal{A}$ is a simple A-Book and $\mathcal{T}$ is a simple T-Book, the $\mathcal{K}$ is a* simple aleatoric knowledge base*, and if $\mathcal{T}$ is also acyclic $\mathcal{K}$ is an* acyclic simple knowledge base*.*

The following lemma is a useful simplification.

Lemma 1. *Every aleatoric knowledge base $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ is equivalent to a simple aleatoric knowledge base, $\mathcal{K}'$.*

See [8] (Lemma ???) for proof.

Theorem 2. *Given a simple aleatoric knowledge base $\mathcal{K} = (\mathcal{A}, \mathcal{T})$ where $\mathcal{T}$ is acyclic, it is possible to determine if $\mathcal{K}$ is consistent with complexity PSPACE.*

The process for the satisfiability theorem is to build a system of polynomial equalities and inequalities corresponding to the axioms in $\mathcal{K}$. This system of constraints is satisfiable if and only if $\mathcal{K}$ is satisfiable. The number of variables, inequalities and equalities in the system is polynomial in the size of $\mathcal{K}$, so determining if the system is satisfiable reduces to $\exists\mathbb{R}$ (the satisfiability of existentially quantified polynomial equations) which is in PSPACE [2]. See [8] (Theorem 4.8) for a full proof. The case for non-acyclic T-Books is left to future work.

3 Expressivity

From Table 1 it can be seen that aleatoric description logic generalises $\mathcal{ALC}$, in the sense that $\mathcal{ALC}$ can be mapped to the $0-1$ fragment of ADL. However, this mapping overlooks the probabilistic aspect of ADL and how this relates to uncertainty in description logics.

The semantic interpretation of aleatoric description logic can be seen as interpreting $\mathcal{ALC}$ over a probability space of interpretations where all roles are functions. The aleatoric belief models of ADL describe a probability space of these simple descriptions, and the semantics of ADL recursively define the likelihood of a formula holding in models sampled from this probability space.

A probability space [13] is a tuple $(\Omega, \mathcal{F}, \mathcal{P})$, where Ω is a set, $\mathcal{F}$ is a σ-algebra over Ω (the events), and $\mathcal{P}$ is a probability measure on $\mathcal{F}$ that is countably additive.

Theorem 3. *Given a formula α of $\mathcal{ALC}$ and an aleatoric belief model $\mathcal{B}_i$, there exists: a formula α^* that is logically equivalent to α in $\mathcal{ALC}$; and probability space $(\Omega^{\mathcal{B}_i}, \mathcal{F}, \mathcal{P}^{\mathcal{B}_i})$ where: $\Omega^{\mathcal{B}_i}$ is a set of functional $\mathcal{ALC}$ models; $\mathcal{F}$ is an algebra over $\Omega^{\mathcal{B}_i}$ consisting of an element $\hat{\beta}$ for every $\mathcal{ALC}$ formula β; and $\mathcal{P}^{\mathcal{B}_i}$ is a probability measure on $\mathcal{F}$ derived from $\mathcal{B}_i$. This probability space is such that $\mathcal{P}^{\mathcal{B}_i}(\hat{\alpha}) = \mathcal{B}_i(\alpha^*)$.*

The proof and necessary constructions can be found in [8] Section 5. This result gives a foundation for the semantics of ADL, establishing a correspondence between the probabilistic operations of ADL, and the deterministic operations of $\mathcal{ALC}$ applied over a probability space of interpretations. This way, ADL represents reasoning where an agent has in mind a probability space of possible interpretations for the state of the world. By observing the world, the agent is able to refine this probability space, and *learn* a better representation for their beliefs.

4 Learning

An aleatoric belief model describes an agent's beliefs and prior assumptions and the agent may update these beliefs based on observations, via Bayesian conditioning. This section will introduce two learning mechanisms, *role learning* and *concept learning* whereby an agent may update the distribution of individuals fulfilling a role, and also update the aleatoric probabilities associated with an atomic concept at an individual. These mechanisms are unique to ADL and provides a compelling advantage over alternative probabilistic description logics [4, 17, 9, 23, 20].

4.1 Role learning

Role learning refines the probability distribution associated with a role ρ. For a pointed aleatoric belief model, $\mathcal{B}_i = (I, r, \ell, i)$, for every $j \in I$, $\rho(i, j)$ is the prior probability that j fulfils the role of ρ for i. Given an observation which is an ADL formula of the form $[\rho]\,(\alpha\,|\,\top)$, $\mathcal{B}_j(\alpha)$ is the probability of this observation holding, given j fulfils the role of ρ for i. Via Bayes' rule, it follows that the probability of j fulfilling the role of ρ for i, *given the observation* is:

$$\rho'(i,j) = \frac{\rho(i,j)\cdot\mathcal{B}_j(\alpha)}{\mathcal{B}_i([\rho]\,(\alpha\,|\,\top))}$$

(the prior probability of j is multiplied by the probability of α given j, divided by the probability of α).

Definition 8. *Let $\mathcal{B}_i = (I, r, \ell, i)$ be an aleatoric belief model, and $\phi = [\rho]\,(\alpha\,|\,\beta)$ an observation, made at i. The ϕ-*update *of $\mathcal{B}_i$ is the aleatoric belief model $\mathcal{B}_i^{\phi} = (I, r^{i,\alpha}, \ell, i)$, where for all $\rho' \neq \rho$ and $j \neq i$, $r^{i,\phi}(\rho', j) = r(\rho, j)$ and for all $j \in I$*

$$r^{i,\phi}(\rho,i)(j) = \frac{\rho(i,j)\cdot\mathcal{B}_j(\alpha)}{\mathcal{B}_i([\rho]\,(\alpha\,|\,\beta))}.$$

Thus an agent with an *aleatoric* model of the world may update their *epistemic* uncertainty of the distribution of roles, via Bayesian conditioning. The ϕ-update of $\mathcal{B}_i$ is the agent's posterior model of the world.

Given the example in Subsection 2.3, suppose that Hector's belief model is $\mathcal{B} = (I, r, \ell, i)$, and Hector is informed that the contact has tested positive for the virus. Hector is also informed that the test used has a 10% false positive rate, so Hector's belief model now includes an atomic concept *FP* that is 0.1 everywhere. Let $\phi = [c]\,((FP?\top{:}V)\,|\,\top)$ and then the ϕ-update of $\mathcal{B}_{\mathtt{H}_0}$ is computed by:

$$r^{\mathtt{H}_0,\phi}(c,\mathtt{H}_0)(j) = \frac{c(i,j)\cdot(0.1 + 0.9\cdot\mathcal{B}_j(V)}{\mathcal{B}_{\mathtt{H}_0}([c]\,((FP?\top{:}V)\,|\,\top))}.$$

Substituting in the values from Table 2, Hector is able to discount the possible individuals without a virus and condition the distribution for *contact* accordingly. The ϕ-update of $\mathcal{B}_{\mathtt{H}_0}$ is represented in Figure 2.

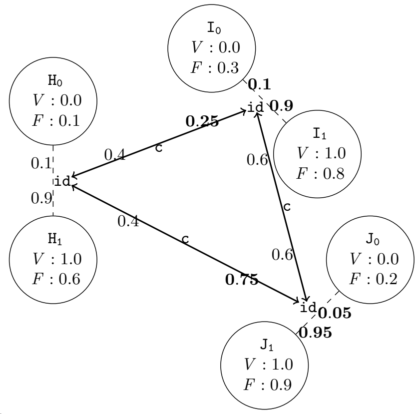


Fig. 2: The ϕ-update of the aleatoric belief model in Figure 1, after Hector is told a contact has tested positive for the virus. The updated values are bold.

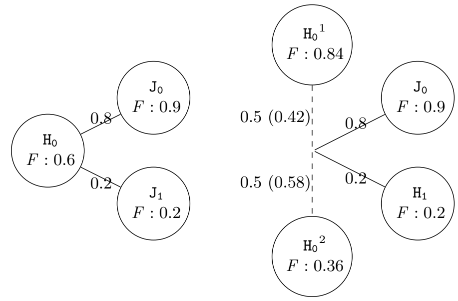


Fig. 3: Concept learning applied to the aleatoric belief model in Figure 1, where Hector applies the belief that he would only have a fever if and only if a contact had a fever. The model on the left is the F-extension, and the probabilities in brackets are the values after role learning has been applied to `id`.

4.2 Concept learning

Role learning is a natural application of Bayes' law since the learning is applied to a probability distribution of possible individuals. However, the probabilities of atomic concepts are modelled as dice, and hence independent of all other variables beyond the possible individual. This means we gain no additional information from applying Bayes' law. If it was possible to observe atomic concepts directly (and often) it would be simple to refine a statistical model of the probabilities. Observations in ADL are complex formulas, so it is preferable to find a more general solution.

Concept learning addresses these issues by introducing new possible individuals in such a way that they do not affect any expected values for named individuals but with variations in the aleatoric probability of concepts, which may then be learnt via role learning, given arbitrary observations. The details of such a construction are given in [8], Section 6.

In the example of Subsection 2.3, suppose that the assessment that the likelihood of Hector having a fever is to be reassessed, based on the observation (or possibly erroneous belief) that Hector would have a fever if and only if Hector's contact had a fever. A new world $\mathtt{H_0}$ is replaced by $\mathtt{H_0}^1$ and $\mathtt{H_0}^2$ where $\mathtt{H_0}^1(F) = 2\mathtt{H_0}(F) - \mathtt{H_0}(f)^2 = 0.84$, and $\mathtt{H_0}^2(F) = \mathtt{H_0}(F)^2 = 0.36$. The probabilities are then updated via role learning over `id`, given the observation $\phi = E_{\mathtt{id}}(F?E_cF{:}E_c\neg F)$, where the relevant fragment of the aleatoric belief model is shown in Figure 3. Note, the model has been revised to make the example clearer. Aggregating $\mathtt{H_0}^1$ and $\mathtt{H_0}^2$ back into a single node by taking weighted sums of the likelihoods gives the updated probability of F at $\mathtt{H_0}$ to be 0.56.

5 Related work

There is a substantial amount of work on logics for reasoning about uncertainty [10], including [15, 14, 24], and going back to the works of Ramsey [21], Carnap [3] and de Finetti [5].

Markov Logic Networks [22] (generalising Bayesian networks and Markov networks) address a similar problem of providing a logical interface to machine learning methods. These approaches attach a probabilistic interpretation to formulas in a fragment of first order logic, rather than providing a probabilistic variation of first order logic operators.

There is some commonality in purpose with probabilistic logic programming [16, 6], although the concepts are constrained to be Horn clauses, where atomic formula are mutually independent.

There is a growing body of work addressing the need for probabilistic reasoning in knowledge bases. In [11], an inductive reasoning approach is applied to include probabilities with rules; in [9], a subjective Bayesian approach is proposed to describe the probabilities associated with a concept or role holding; and Lukasiewicz and Straccia [17] have proposed a method to include vagueness (or fuzzy concepts [25]) in descriptions logics. Probabilistic extensions of description logics have also been proposed by Rigguzzi et al [23] and Pozzato [20]. These approaches extend knowledge bases to include probabilistic assertions and axioms, and provide an extended syntax for querying probability thresholds. Some work on learning parameters and structure of knowledge bases via probabilistic description logics has been done, including Ceylan and Penaloza [4], who have proposed a Bayesian Description Logic that combines a basic description logic framework with Bayesian networks [19] for representing uncertainty about facts, and Ochoa Luna et al [18] who applied statistical methods to estimate the most likely configuration of a knowledge base.

These approaches are very different to the work presented here, as probabilities are not propagated through the roles, and they do not permit learning based on the observation of complex propositions.

6 Conclusion

This paper has introduced a novel approach for representing uncertain knowledge and beliefs. Generalising the description logic $\mathcal{ALC}$, the aleatoric description logic is able to represent complex concepts as independent aleatoric events. The events are contingent on *possible individuals* so they give a subjective Bayesian interpretation of knowledge bases. This paper has also given computational reasoning methods for aleatoric knowledge bases, and shown how aleatoric description logic corresponds to a probability space of functional $\mathcal{ALC}$ models. The aleatoric concepts and roles enable a simple learning framework where agents are able to update their beliefs based on the observations of complex propositions.

Future work will examine the complexity of the satisfiability problem for non-acyclic T-Books, and investigate implementing a reasoning system for aleatoric description logics.

References

1. Baader, F., Calvanese, D., McGuinness, D., Patel-Schneider, P., Nardi, D., et al.: The description logic handbook: Theory, implementation and applications. Cambridge university press (2003)
2. Canny, J.: Some algebraic and geometric computations in PSPACE. In: STOC '88 (1988)
3. Carnap, R.: On Inductive Logic. Philosophy of science **12**(2), 72–97 (1945)
4. Ceylan, I.I., Penaloza, R.: The Bayesian Description Logic $\mathcal{BEL}$. In: IJCAR. pp. 480–494. Springer (2014)
5. De Finetti, B.: Theory of Probability: A critical introductory treatment. John Wiley & Sons (1970)
6. De Raedt, L., Kimmig, A., Toivonen, H.: Problog: A probabilistic prolog and its application in link discovery. In: IJCAI. vol. 7, pp. 2462–2467. Hyderabad (2007)
7. French, T., Gozzard, A., Reynolds, M.: A Modal Aleatoric Calculus for Probabilistic Reasoning. In: ICLA. pp. 52–63. Springer (2019)
8. French, T., Smoker, T.: An aleatoric description logic for probabilistic reasoning (long version) (2021), http://arxiv.org/abs/2108.13036
9. Gutiérrez-Basulto, V., Jung, J.C., Lutz, C., Schröder, L.: Probabilistic Description Logics for Subjective Uncertainty. Journal of Artificial Intelligence Research **58**, 1–66 (2017)
10. Halpern, J.: Reasoning about Uncertainty. MIT Press, Cambridge MA (2003)
11. Heinsohn, J.: Probabilistic Description Logics. In: Uncertainty Proceedings 1994, pp. 311–318. Elsevier (1994)
12. Jøsang, A.: Subjective logic. Springer (2016)
13. Kolmogorov, A.N., Bharucha-Reid, A.T.: Foundations of the Theory of Probability: Second English Edition. Courier Dover Publications (2018)
14. Kooi, B.P.: Probabilistic Dynamic Epistemic Logic. Journal of Logic, Language and Information **12**(4), 381–408 (2003)
15. Kozen, D.: A Probabilistic PDL. Journal of Computer and System Sciences **30**(2), 162–178 (1985)
16. Lukasiewicz, T.: Probabilistic logic programming. In: ECAI. pp. 388–392 (1998)
17. Lukasiewicz, T., Straccia, U.: Managing Uncertainty and Vagueness in Description Logics for the Semantic Web. Web Semantics: Science, Services and Agents on the World Wide Web **6**(4), 291–308 (2008)
18. Ochoa-Luna, J.E., Revoredo, K., Cozman, F.G.: Learning probabilistic description logics: A framework and algorithms. In: Mexican International Conference on Artificial Intelligence. pp. 28–39. Springer (2011)
19. Pearl, J.: Causality: Models, Reasoning, and Inference. Econometric Theory **19**(675-685), 46 (2003)
20. Pozzato, G.L.: Typicalities and probabilities of exceptions in nonmotonic description logics. International Journal of Approximate Reasoning **107**, 81–100 (2019)
21. Ramsey, F.P.: Truth and probability (1926). The Foundations of Mathematics and other Logical Essays pp. 156–198 (1931)
22. Richardson, M., Domingos, P.: Markov logic networks. Machine learning **62**(1-2), 107–136 (2006)
23. Riguzzi, F., Bellodi, E., Lamma, E., Zese, R.: Probabilistic description logics under the distribution semantics. Semantic Web **6**(5), 477–501 (2015)
24. Van Benthem, J., Gerbrandy, J., Kooi, B.: Dynamic Update with Probabilities. Studia Logica **93**(1), 67 (2009)
25. Zadeh, L.A.: Fuzzy Sets. In: Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh, pp. 394–432. World Scientific (1996)