# Aleatoric Predicates: Reasoning about Marbles

Extended Abstract

Tim French
School of Physics, Mathematics and Computing, The University of Western Australia
Perth, Western Australia
tim.french@uwa.edu.au

## ABSTRACT

Aleatoric Logic is the logic of dice, where Boolean propositions are replaced by independent probabilistic events. In a first order extension of this notion, Aleatoric predicates are applied to domain elements selected via independent probabilistic events. An analogy for this is the classic *marbles in an urn* problem, where we might ask the probability of drawing three marbles of the same colour from an urn, or drawing only black marbles from an urn until a red marble is drawn. This paper formalises a syntax and semantics for propositions built from aleatoric predicates, and discusses how these predicates give a representation of an agent's beliefs that come through experience.

## KEYWORDS

Aleatoric Logic; Probabilistic Reasoning; Learning Agents

**ACM Reference Format:**
Tim French. 2024. Aleatoric Predicates: Reasoning about Marbles: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

## 1 INTRODUCTION

Aleatoric logic [5, 6] is a generalisation of propositional logic, where the Boolean values (true and false) are replaced with probabilities, in a similar way to fuzzy logics [15]. However, unlike fuzzy logics, propositions are treated as events (coin flips, or rolls of a dice), so we can define complex and conditional propositions describing sets of events, such as *rolling the same number in a die three times in a row*. Here we extend that notion to include predicates defined over some probability space. The analogy is drawing marbles from an urn, where the marbles have some colour or label. For example, we may suppose that we have an urn containing marbles labelled with positive integers. We can draw a marble until we find one whose label is prime, note its label and replace it, and draw a second. Given such a process, what is the probability that the second marble has a label less than the first? If we knew all the marbles in the urn, we could easily calculate this, but without full knowledge an agent must defer to belief. An agent may have experience drawing marbles from this urn, and had only ever seen large prime labels, so may suppose the chance is quite high. A different agent with no experience may suppose that the primes become increasingly sparse at high values so may reason the chance is small. Importantly, as the outcome of a predicate may be contingent on the value of some marble drawn earlier, the probabilities of propositions are no longer independent, and allow us to express complex dependencies between predicates.

There has been a considerable number of works that have considered probabilistic semantics and logics. Early work includes Kolmogorov [11], Ramsey [13] and de Finetti [2], who produced axioms for reasoning about probabilities of events, and [7, 9, 10, 12] give an overview of probabilistic reasoning. The approach presented here has similarities to fuzzy logic [15], but with a focus on dynamic belief [1, 3, 8].

## 2 ALEATORIC PREDICATES

For the syntax we assume a set of *domain variables* $\mathcal{V}$, a set of *propositional atoms* $\mathcal{A}$, and a set of *predicates* $\mathcal{P}$ where each predicate $P \in \mathcal{P}$ has an arity $\#P$ which is some positive integer.

Definition 1. *The syntax for* $\mathcal{L}$*, the* language of aleatoric propositions*, is given by the Backus-Naur form:*

$$\alpha ::= \top \mid X \mid P(x_1, ..., x_{\#P}) \mid \neg\alpha \mid (\alpha \; ? \; \alpha : \alpha) \mid \mathbb{E}x.\alpha \mid \mathbb{F}X.\alpha$$

*where* $X \in \mathcal{A}$*,* $P \in \mathcal{P}$*,* $x \in \mathcal{V}$*, and* $X$ is linear in $\alpha$*: an atomic proposition* $X \in \mathcal{A}$ is linear in $\alpha$ *if and only if for every sub-formula* $(\beta \; ? \; \gamma_1 : \gamma_2)$ *of* $\alpha$*, there is no occurrence of* $X$ *appearing in* $\beta$*.*

We write $P(\overline{x})$ as an abbreviation for $P(x_1, ..., x_{\#P})$. Each formula $\alpha \in \mathcal{L}$ can be thought of as a proposition describing a set of events. In this setting a brief description of these operators is as follows:

$\top$ is the *true* event, so it is always true;

$X$, where $X \in \mathcal{A}$ is an atomic proposition, which describes some independent random event, like a coin landing heads;

$P(\overline{x})$ is a predicate over the domain variables $x_1, ..., x_{\#P}$, where he interpretation is fixed, so given $x_1, ...x_{\#P}$ it is either always true or always false;

$\neg\alpha$, describes the *failure* of an event to occur, so the event is explicitly tested for, and that test fails;

$(\alpha \; ? \; \beta : \gamma)$ describes a *conditional* event where the event described by $\alpha$ is tested (or sampled) and if it occurs, then an event described by $\beta$ occurs, but if an alpha event does not occur, then an event described by $\gamma$ occurs;

$\mathbb{E}x.\alpha$ is the *expectation* operator, and expresses the likelihood $\alpha$ will be true when $x$ is drawn randomly from the domain;

$\mathbb{F}X.\alpha$ is the *fixed point* operator, and it describes an event with probability $p$ such that if the event corresponding to the atom $X$ had likelihood $p$, so would $\alpha$. (See [4] for a discussion of this operator).

For an example of the syntax, consider the example from the introduction, where we have an urn of marbles labelled by integers,



*Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.* © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

and we sample from the urn until we have drawn a marble, $x$ with a prime label, and then we draw a second marble, $y$, whose label is less than the label of $x$. This can be represented as

$$\mathbb{F}X.\mathbb{E}x.(prime(x) \;?\; \mathbb{E}y.y < x : X).$$

In this case $X$ is a fixed point variable, so in the event $x$ is sampled and is not prime, the whole formula is substituted for $X$, which effectively requires $x$ to be resampled.

We interpret aleatoric predicate logic over a first order domain, where domain elements are sampled from a probability space and propositional atoms are assigned to probabilities.

Definition 2. *An* aleatoric interpretation *is given by the tuple: $\mathcal{I} = (\mathcal{S}, \Sigma, \mu, \chi, \nu)$ where:*

- *$(\mathcal{S}, \Sigma, \mu)$ is a probability space, consisting of a domain $\mathcal{S}$, a $\sigma$-algebra over $\mathcal{S}$, $\Sigma$, and a probability measure over $\Sigma$, $\mu$.*
- *$\chi \in [0,1]^{\mathcal{A}}$ assigns a probability to each atomic proposition.*
- *$\nu \in \wp(\mathcal{S}^*)^{\mathcal{P}}$ assigns each predicate $P \in \mathcal{P}$ to a set of tuples over the $\mathcal{S}$, where all tuples in $\nu(P)$ have length $\#P$ (i.e. for all $w \in \nu(P)$, $|w| = \#P$).*

We note the assignment $\nu$ gives a first order interpretation of the set of predicates, $\mathcal{P}$, over the domain $\mathcal{S}$. That is, given $P \in \mathcal{P}$ and some assignment of variables to domain elements $a \in \mathcal{S}^{\mathcal{V}}$, the predicate $P(\overline{x})$ holds if $(a(x_1), \dots, a(x_{\#P})) \in \nu(P)$, and $\mathcal{S}^*$ is the set of finite words over the alphabet $\mathcal{S}$.

Definition 3. *Given an interpretation, $\mathcal{I} = (\mathcal{S}, \Sigma, \mu, \chi, \nu)$ and some $\alpha \in \mathcal{L}$, the* likelihood of $\alpha$ in $\mathcal{I}$ *is a function $\alpha^{\mathcal{I}} : \mathcal{S}^{\mathcal{V}} \longrightarrow [0,1]$, specified inductively as follows. Given some $a \in \mathcal{S}^{\mathcal{V}}$:*

$$\begin{array}{rcl}
\top^{\mathcal{I}} & = & 1 \\
X^{\mathcal{I}} & = & \chi(X) \\
(P(\overline{x}))^{\mathcal{I}}(a) & = & 1 \text{ if } (a(x_1), ..., a(x_{\#P})) \in \nu(P), \\
 & & \text{and } 0 \text{ otherwise} \\
(\alpha \;?\; \beta : \gamma)^{\mathcal{I}} & = & \alpha^{\mathcal{I}} \cdot \beta^{\mathcal{I}} + (1 - \alpha^{\mathcal{I}}) \cdot \gamma^{\mathcal{I}} \\
(\neg\alpha)^{\mathcal{I}} & = & 1 - \alpha^{\mathcal{I}} \\
(\mathbb{E}x.\alpha)^{\mathcal{I}} & = & \int_{\mathcal{S}} \alpha^{\mathcal{I}} \, d\mu(x) \\
(\mathbb{F}X.\alpha)^{\mathcal{I}} & = & \begin{cases} 1/2 \text{ if } (\alpha^{\mathcal{I}})^{X \leftarrow 1/2} = 1/2, \\ p \text{ if } (\alpha^{\mathcal{I}})^{X \leftarrow p} = p \text{ and } p \neq 1/2 \end{cases}
\end{array}$$

*We call $\alpha^{\mathcal{I}}$ the* descriptive interpretation of $\alpha$. *Note only the interpretation of predicates is dependent on the assignment $a \in \mathcal{S}^{\mathcal{V}}$. Where the assignment $a \in \mathcal{S}^{\mathcal{V}}$ is clear from context, we write $\mathcal{I}(\alpha)$ in place of $\alpha^{\mathcal{I}}(a)$.*

## 3 ALEATORIC LEARNING THEORY

An interpretation (Definition 2) represents an agent's subjective beliefs, so the evaluations of propositions as events may be thought of as imagined events, or simulations derived from an agent's beliefs (see [14]). These beliefs are built through experience, where an agent observes some external event, and this observation is then incorporated into their experience, and affects their beliefs. The observation is a query: $\alpha(x)$, by which we suppose that a single element is sampled from the domain, and $\alpha(x)$ is tested, so it either happened, or it did not. This is a single bit of information and is akin to looking out of a window to see if it is raining. Aleatoric learning occurs via the probability measure $\mu$ in the interpretation $\mathcal{I} = (\mathcal{S}, \Sigma, \mu, \chi, \nu)$, and all other parts of the interpretation are fixed.

To give the basic principle, let us return to the metaphor of the urn of labelled marbles. The observations come from an environment, which we suppose behaves just like an (external) urn of marbles, while the agent (the observer) has their own set of beliefs, which we consider to be an internal urn of marbles. The agent would like to condition their beliefs given their observations

Suppose that there is a very large number of marbles in the agent's internal urn, and the agent is able to draw a marble and test whether $\alpha(x)$ holds for that marble. The agent can also observe the environment which gives them one bit of information at a time: whether $\alpha(x)$ was true for a marble randomly sampled from the external urn. The learning operation is as follows.

**1.** Take two empty urns, and sample a large number of marbles from the full urn, in such a way that marbles satisfying $\alpha(x)$ are more likely to be in the left urn, (it is *$\alpha$-supportive*) while marbles not satisfying $\alpha(x)$ are more likely to be in the right urn (it is *$\alpha$-sceptical*).

**2.** Suppose a process where the agent flips a coin and: if it lands heads, they draw a single marble from the left urn and test whether $\alpha$ is true for that marble; and if it lands tails, they draw a single marble from the right urn and test whether $\alpha$ is true for that marble.

**3.** The agent makes an observation of the environment and applies Bayesian conditioning to estimate the probability $p_\ell$ that the marble came from the left urn (so the chance the marble came from the right urn is $1 - p_\ell$).

**4.** To update the agent's beliefs they empty the marbles from the original urn, and now fill it with marbles taken from the left and right urn, where there is a $p_\ell$ chance that each next marble is taken from the left urn.

The original urn now contains a subset of marbles which represents the agent's learnt interpretation $\mathcal{I}^\alpha$ given the observation $\alpha$. Now, given an observation $\alpha(x)$ we can *syntactically* evaluate a formula, $\beta$ in the two subdomains weighted by how likely $\alpha(x)$ is in each subdomain.

Definition 4. *Given an observation of $\alpha(x)$, the conditioning of $\beta$ by $\alpha(x)$ is the proposition:*

$$\beta^\alpha = \beta[\mathbb{E}x.\gamma(x) \backslash \mathbb{F}X.(1/2 \;?\; (\alpha^\ell \;?\; \gamma^\ell : X) : (\alpha^r \;?\; \gamma^r : X))]$$

*where for any proposition $\delta$, $\delta^r = \delta[\mathbb{E}x.\gamma \backslash \mathbb{E}x.(\alpha \;?\; \mathbb{E}x.\gamma : \gamma)]$ and $\delta^\ell = \delta[\mathbb{E}x.\gamma \backslash \mathbb{E}x.(\alpha \;?\; \gamma : \mathbb{E}x.\gamma)]$.*

In this definition $\beta^r$ favours the part of the domain where $\alpha$ is less likely by, whenever sampling some $x$ to evaluate $\gamma$, first testing $\alpha$, and if $\alpha$ is true, requiring that $x$ should be resampled before testing $\gamma$. Similarly $\beta^\ell$ favour the part of the domain where $\alpha$ is more likely, by first testing $\alpha$ and only resampling if $\alpha$ is false. In this way we simulate conditioning the two subdomains on the observation of $\alpha$, so given any aleatoric proposition $\beta$ there is a computable aleatoric proposition $\alpha^\beta$ such that the following theorem holds.

Theorem 5. *Given some interpretation $\mathcal{I}$, an aleatoric proposition $\alpha$ with one free domain variable $x$, and an aleatoric proposition $\beta$: $(\beta^\alpha)^{\mathcal{I}} = (\beta)^{\mathcal{I}^\alpha}$*

## REFERENCES


[1] Alexandru Baltag, Soroush Rafiee Rad, and Sonja Smets. 2021. Tracking probabilistic truths: a logic for statistical learning. *Synthese* 199 (2021), 9041–9087.

[2] Bruno De Finetti. 1970. *Theory of probability: a critical introductory treatment.* John Wiley & Sons.

[3] H. van Ditmarsch, W. van der Hoek, and B. Kooi. 2007. *Dynamic Epistemic Logic.* Synthese Library, Vol. 337. Springer.

[4] Tim French. 2023. Aleatoric Propositions: Reasoning About Coins. In *Logic, Language, Information, and Computation: 29th International Workshop, WoLLIC 2023, Halifax, NS, Canada, July 11?14, 2023, Proceedings*, Helle Hvid Hansen, Andre Scedrov, and Ruy J. G. B. de Queiroz (Eds.). Springer Nature Switzerland, 227–243.

[5] Tim French, Andrew Gozzard, and Mark Reynolds. 2019. Dynamic aleatoric reasoning in games of bluffing and chance. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. 1964–1966.

[6] Tim French, Andrew Gozzard, and Mark Reynolds. 2019. A modal aleatoric calculus for probabilistic reasoning. In *Logic and Its Applications: 8th Indian Conference, ICLA 2019, Delhi, India, March 1-5, 2019, Proceedings 8*. Springer, 52–63.

[7] Peter Gärdenfors. 1975. Qualitative probability as an intensional logic. *Journal of Philosophical Logic* (1975), 171–185.

[8] Peter Gärdenfors. 2003. *Belief revision.* Number 29. Cambridge University Press.

[9] J.Y. Halpern. 2003. *Reasoning about Uncertainty.* MIT Press, Cambridge MA.

[10] Hykel Hosni and Jürgen Landes. 2023. Logical perspectives on the foundations of probability. *Open Mathematics* 21, 1 (2023), 20220598.

[11] Andrei Nikolaevich Kolmogorov. 1963. The theory of probability. *Mathematics, Its Content, Methods, and Meaning* 2 (1963), 110–118.

[12] Dexter Kozen. 1985. A probabilistic PDL. *J. Comput. System Sci.* 30, 2 (1985), 162–178.

[13] Frank Plumpton Ramsey. 1925. The foundations of mathematics. (1925).

[14] Timothy Williamson. 2016. Knowing by Imagining. In *Knowledge Through Imagination*. Oxford University Press.

[15] Lotfi A Zadeh. 1996. Fuzzy sets. In *Fuzzy Sets, Fuzzy Logic, And Fuzzy Systems: Selected Papers by Lotfi A Zadeh*. World Scientific, 394–432.