

A Modal Characterization Theorem for a Probabilistic Fuzzy Description Logic

Paul Wild¹, Lutz Schröder¹, Dirk Pattinson² and Barbara König³

¹Friedrich-Alexander-Universität Erlangen-Nürnberg

²Australian National University, Canberra

³Universität Duisburg-Essen

{paul.wild,lutz.schroeder}@fau.de, dirk.pattinson@anu.edu.au, barbara.koenig@uni-due.de

Abstract

The fuzzy modality *probably* is interpreted over probabilistic type spaces by taking expected truth values. The arising probabilistic fuzzy description logic is invariant under probabilistic bisimilarity; more informatively, it is non-expansive wrt. a suitable notion of behavioural distance. In the present paper, we provide a characterization of the expressive power of this logic based on this observation: We prove a probabilistic analogue of the classical van Benthem theorem, which states that modal logic is precisely the bisimulation-invariant fragment of first-order logic. Specifically, we show that every formula in probabilistic fuzzy first-order logic that is non-expansive wrt. behavioural distance can be approximated by concepts of bounded rank in probabilistic fuzzy description logic.

1 Introduction

In the representation of uncertain knowledge, one will often wish to avoid mention of exact numerical probabilities, e.g. when these are not precisely known or not relevant to the representation task at hand – as a typical example, a medical practitioner will rarely name a numerical threshold for the likelihood of a diagnosis, and instead qualify the diagnosis as, say, ‘suspected’ or ‘probable’. This has led to efforts aimed at formalizing a modality *probably*, used alternatively to modalities ‘with probability at least p ’ [Larsen and Skou, 1991; Heifetz and Mongin, 2001]. Such a formalization may be approached in a two-valued setting via qualitative axiomatizations of likelihood [Burgess, 1969; Halpern and Rabin, 1987] or via threshold probabilities [Hamblin, 1959; Herzig, 2003]. In a fuzzy setting, ‘probably’ leads a natural life as a fuzzy modality P , whose truth value just increases as its argument becomes more probable (this modality thus connects the otherwise well-distinguished worlds of fuzziness and probability [Lukasiewicz and Straccia, 2008]). The semantics of this operator, first defined by Zadeh [1968], interprets $P\phi$ as the expected truth value of ϕ . It appears in various fuzzy propositional [Hájek, 2007; Flaminio and Godo, 2007], modal [Desharnais *et al.*, 2004; van Breugel and Worrell, 2005], fixpoint [Kozen, 1985; Huth

and Kwiatkowska, 1997], and description logics [Schröder and Pattinson, 2011].

In the present paper, we pin down the exact expressiveness of the basic description logic of *probably*, which we briefly refer to as *probabilistic fuzzy \mathcal{ALC}* or $\mathcal{ALC}(P)$, within a natural ambient probabilistic fuzzy first-order logic $\text{FO}(P)$, by providing a *modal characterization theorem*. The prototype of such characterization theorems is *van Benthem’s theorem* [1976], which states that (classical) modal logic is precisely the bisimulation-invariant fragment of first-order logic. It has been noted that in systems with numerical values, *behavioural pseudometrics* offer a more fine-grained measure of equivalence than two-valued bisimilarity [Giacalone *et al.*, 1990; Desharnais *et al.*, 2004; van Breugel and Worrell, 2005; Desharnais *et al.*, 2008; Baldan *et al.*, 2014]. When propositional connectives are equipped with Zadeh semantics, $\mathcal{ALC}(P)$ is *non-expansive* wrt. behavioural distance; we continue to refer to this property as *bisimulation invariance*. In previous work [Wild *et al.*, 2018] we have shown that *relational fuzzy modal logic* is the bisimulation-invariant fragment of fuzzy FOL, more precisely that every bisimulation-invariant fuzzy FO formula can be approximated by fuzzy modal formulae *of bounded rank*. The bound on the rank is key; without it, the statement turns into a form of the (much simpler) Hennessy-Milner theorem [Hennessy and Milner, 1985] (which classically states that non-bisimilar states in finitely branching systems can be distinguished by modal formulae), and indeed does not need to assume FO definability of the given bisimulation-invariant property [van Breugel and Worrell, 2005]. Here, we establish a corresponding result for the rather more involved probabilistic setting: We show that *every bisimulation-invariant formula in probabilistic fuzzy FOL can be approximated in bounded rank in probabilistic fuzzy \mathcal{ALC}* . This means not only that, up to approximation, $\mathcal{ALC}(P)$ is as powerful as $\text{FO}(P)$ on bisimulation-invariant properties, but also that $\mathcal{ALC}(P)$ provides effective syntax for bisimulation-invariant $\text{FO}(P)$, which $\text{FO}(P)$ itself does not.

A full version is available at <http://arxiv.org/abs/1906.00784>.

Related Work There is widespread interest in modal characterization theorems in many areas, including concurrency [Janin and Walukiewicz, 1995] as well as AI and database theory [Sturm and Wolter, 2001; Figueira *et al.*, 2015; Wild and Schröder, 2017; Wild *et al.*, 2018]. The

overall structure of our proof builds partly on that of our modal characterization theorem for relational fuzzy modal logic [Wild *et al.*, 2018] (in turn based ultimately on a strategy due to Otto [2004]) but deals with a much more involved logic, which instead of just the lattice structure of the unit interval involves its full arithmetic structure, via the use of probabilities and expected values, necessitating, e.g., the use of Kantorovich-Rubinstein duality. Notable contributions of our proof include new forms of probabilistic bisimulation games up-to- ϵ (different from games introduced by Desharnais *et al.* [2008], which characterize a different metric) and Ehrenfeucht-Fraïssé games, related to two-valued games considered in the context of topological FOL [Makowsky and Ziegler, 1980]. (For lack of space, we omit discussion of quantitative Hennessy-Milner type results beyond the mentioned result by van Breugel and Worrell [2005].)

FO(P) may be seen as a fuzzy variant of Halpern’s [1990] type-1 (i.e. statistical) two-valued probabilistic FOL, and uses a syntax related to coalgebraic predicate logic [Litak *et al.*, 2018]. Van-Benthem style theorems for two-valued coalgebraic modal logic [Schröder *et al.*, 2017] instantiate to two-valued probabilistic modal logic, then establishing expressibility of bisimulation-invariant probabilistic FO formulae by probabilistic modal formulae with infinite conjunction but of bounded rank, in an apparent analogy to bounded-rank approximation in the fuzzy setting.

2 Fuzzy Probabilistic Logics

We proceed to introduce the logics featuring in our main result. We fix (w.l.o.g., finite) sets N_C of *atomic concepts* and N_R of *roles*; *concepts* C, D of *quantitative probabilistic ALC* ($ALC(P)$) are defined by the grammar

$$C, D ::= q \mid A \mid C \ominus q \mid \neg C \mid C \sqcap D \mid \mathbf{P} r. C$$

where $q \in \mathbb{Q} \cap [0, 1]$, $A \in N_C$ and $r \in N_R$. The intended reading of \mathbf{P} is ‘probably’; we give examples below. Slightly deviating from standard practice, we define the *rank* $\text{rk}(C)$ of a concept C as the maximal nesting depth of the \mathbf{P} and *atomic concepts* in C ; e.g. $\text{rk}((\mathbf{P} r. \mathbf{P} s. A) \sqcap (\mathbf{P} r. B)) = 3$. We denote the set of all concepts of rank at most n by $ALC(P)_n$.

Concepts are interpreted over probabilistic structures to which we neutrally refer as *interpretations* or, briefly, *models*. We allow infinite models but restrict to discrete probability distributions over successors at each state. A model

$$\mathcal{I} = (\Delta^{\mathcal{I}}, (A^{\mathcal{I}})_{A \in N_C}, (r^{\mathcal{I}})_{r \in N_R})$$

consists of a *domain* $\Delta^{\mathcal{I}}$ of *states* or *individuals*, and interpretations $A^{\mathcal{I}}: \Delta^{\mathcal{I}} \rightarrow [0, 1]$, $r^{\mathcal{I}}: \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow [0, 1]$ of atomic concepts A and roles r such that for each $a \in \Delta^{\mathcal{I}}$, the map

$$r_a: \Delta^{\mathcal{I}} \rightarrow [0, 1], \quad r_a(a') = r^{\mathcal{I}}(a, a')$$

is either zero or a probability mass function on $\Delta^{\mathcal{I}}$, i.e.

$$\sum_{a' \in \Delta^{\mathcal{I}}} r_a(a') \in \{0, 1\}$$

(implying that the *support* $\{a' \in \Delta^{\mathcal{I}} \mid r_a(a') > 0\}$ of r_a is at most countable). We call a state a *r-blocking* if $\sum_{a' \in \Delta^{\mathcal{I}}} r_a(a') = 0$. At non-blocking states a , r_a thus acts

as a probabilistic accessibility relation; we abuse r_a to denote also the probability measure defined by r_a .

The interpretation $C^{\mathcal{I}}: \Delta^{\mathcal{I}} \rightarrow [0, 1]$ of concepts is defined recursively, extending that of atomic concepts, by

$$\begin{aligned} q^{\mathcal{I}}(a) &= q \\ (C \ominus q)^{\mathcal{I}}(a) &= \max(C^{\mathcal{I}}(a) - q, 0) \\ (\neg C)^{\mathcal{I}}(a) &= 1 - C^{\mathcal{I}}(a) \\ (C \sqcap D)^{\mathcal{I}}(a) &= \min(C^{\mathcal{I}}(a), D^{\mathcal{I}}(a)) \\ (\mathbf{P} r. C)^{\mathcal{I}}(a) &= E_{r_a}(C^{\mathcal{I}}) = \sum_{a' \in \Delta^{\mathcal{I}}} r_a(a') \cdot C^{\mathcal{I}}(a') \end{aligned}$$

At non-blocking a , $(\mathbf{P} r. C)^{\mathcal{I}}(a)$ is thus the expected truth value of C for a random r -successor of a . We define disjunction \sqcup as the dual of \sqcap as usual, so \sqcup takes maxima. We use Zadeh semantics for the propositional operators, which will later ensure non-expansiveness wrt. behavioural distance; see additional comments in Section 7.

Up to minor variations, our models correspond to Markov chains or, in an epistemic reading, *type spaces* (e.g. [Heifetz and Mongin, 2001]). The logic $ALC(P)$ was considered (with Łukasiewicz semantics) by Schröder and Pattinson [2011], and resembles van Breugel and Worrell’s quantitative probabilistic modal logic [2005]. E.g., in a reading of $\Delta^{\mathcal{I}}$ as consisting of real-world individuals, the concept

Loud \sqcap P hasSource. (Large \sqcap P hasMood. Angry)

describes noises you hear in your tent at night as being loud and probably coming from the large and probably angry animal whose shadow just crossed the tent roof. (In this view, \mathbf{P} can be usefully combined with crisp or fuzzy relational modalities, using off-the-shelf compositionality mechanisms [Schröder and Pattinson, 2011].) In an epistemic reading where the elements of $\Delta^{\mathcal{I}}$ are possible worlds, and the roles are understood as epistemic agents, the concept

\neg GoodHand \sqcap P player. P opponent. GoodHand

denotes the degree to which player believes she is successfully bluffing by letting opponent overestimate player’s hand.

For readability, we will restrict the technical treatment to a single role r , omitted in the syntax, from now on, noting that covering multiple roles amounts to no more than additional indexing. As the first-order correspondence language of quantitative probabilistic ALC we introduce *quantitative probabilistic first-order logic* (FO(P)), with *formulae* ϕ, ψ, \dots defined by the grammar

$$\begin{aligned} \phi, \psi ::= q \mid A(x) \mid x = y \mid \phi \ominus q \mid \neg \phi \mid \phi \sqcap \psi \mid \exists x. \phi \\ \mid x\mathbf{P}[y : \phi] \quad (q \in \mathbb{Q} \cap [0, 1], A \in N_C) \end{aligned}$$

where x and y range over a fixed countably infinite reservoir of *variables*. The reading of $x\mathbf{P}[y : \phi]$ is the expected truth value of ϕ at a random successor y of x . (In particular, when ϕ is crisp, then $x\mathbf{P}[y : \phi]$ is just the probability of y satisfying ϕ , similar to the weights $w_y(\phi)$ in Halpern’s type-1 probabilistic FOL [1990].) We have the expected notions of free and bound variables, under the additional proviso that y (but not x !) is bound in $x\mathbf{P}[y : \phi]$. The (*quantifier*) *rank* $\text{qr}(\phi)$ of a formula ϕ is the maximal nesting depth of the

variable-binding operators \exists and \mathbf{P} and propositional atoms A in ϕ ; e.g. $\exists x. x\mathbf{P}[y : A(y)]$ has rank 3.

Given a model $\mathcal{I} = (\Delta^{\mathcal{I}}, (A^{\mathcal{I}})_{A \in \mathbf{Nc}}, r^{\mathcal{I}})$ and a vector $\bar{a} = (a_1, \dots, a_n) \in (\Delta^{\mathcal{I}})^n$ of values, the semantics of the logic assigns a truth value $\phi(\bar{a}) \in [0, 1]$ to a formula $\phi(x_1, \dots, x_n)$ with free variables at most x_1, \dots, x_n . We define $\phi(\bar{a})$ recursively by essentially the same clauses as in $\mathcal{ALC}(\mathbf{P})$ for the propositional constructs, and

$$A(x_i)(\bar{a}) = A^{\mathcal{I}}(a_i)$$

$$(\exists x_0. \phi(x_0, x_1, \dots, x_n))(\bar{a}) = \bigvee_{a_0 \in \Delta^{\mathcal{I}}} \phi(a_0, a_1, \dots, a_n)$$

$$(x_i\mathbf{P}[y : \phi(y, x_1, \dots, x_n)])(\bar{a}) = E_{r_{a_i}}(\phi(\cdot, a_1, \dots, a_n))$$

where \bigvee takes suprema. Moreover, equality is two-valued, i.e. $(x_i = x_j)(\bar{a})$ is 1 if $a_i = a_j$, and 0 otherwise.

E.g. the formula $x\mathbf{P}[z : z = y]$ ('the successor of x is probably y ') denotes the access probability from x to y , $x\mathbf{P}[z : z\mathbf{P}[w : w = y]]$ the probability of reaching y from x in two independently distributed steps, and $\exists y. x\mathbf{P}[z : z = y]$ the probability of the most probable successor of x .

We have a *standard translation* ST_x from $\mathcal{ALC}(\mathbf{P})$ into $\mathbf{FO}(\mathbf{P})$, indexed over a variable x naming the current state. Following Litak et al. [2018], we define ST_x recursively by

$$ST_x(A) = A(x)$$

$$ST_x(\mathbf{PC}) = x\mathbf{P}[y : ST_y(C)],$$

and by commutation with all other constructs.

Lemma 2.1. *For every $\mathcal{ALC}(\mathbf{P})$ -concept C and state a , $C(a) = ST_x(C)(a)$.*

ST thus identifies $\mathcal{ALC}(\mathbf{P})$ as a fragment of $\mathbf{FO}(\mathbf{P})$.

3 Behavioural Distances and Games

We next discuss several notions of behavioural distance between states: via fixed point iteration à la Wasserstein/Kantorovich, via games and via the logic. We focus mostly on depth- n distances. Only for one version, we define also the unbounded distance, which will feature in the modal characterization result. We show in Section 4 that at finite depth, all these distances coincide. It has been shown in previous work [Desharnais et al., 2004; van Breugel and Worrell, 2005] that the unbounded-depth distances defined via Kantorovich fixed point iteration and via the logic, respectively, coincide in very similar settings; such results can be seen as probabilistic variants of the Hennessy-Milner theorem.

We recall standard notions on pseudometric spaces:

Definition 3.1 (Pseudometric spaces, non-expansive maps). A (*bounded*) *pseudometric* on a set X is a function $d: X \times X \rightarrow [0, 1]$ such that for $x, y, z \in X$, the following axioms hold: $d(x, x) = 0$ (*reflexivity*), $d(x, y) = d(y, x)$ (*symmetry*), $d(x, z) \leq d(x, y) + d(y, z)$ (*triangle inequality*). If additionally $d(x, y) = 0$ implies $x = y$, then d is a *metric*. A (*pseudo*)*metric space* (X, d) consists of a set X and a (*pseudo*)*metric* d on X .

A map $f: X \rightarrow [0, 1]$ is *non-expansive* wrt. a pseudometric d if $|f(x) - f(y)| \leq d(x, y)$ for all $x, y \in X$. The space of these non-expansive functions, denoted $\text{Pred}(X, d)$, is equipped with the *supremum (pseudo)metric* d_∞ ,

$$d_\infty(f, g) = \|f - g\|_\infty = \bigvee_{x \in X} |f(x) - g(x)|.$$

We denote by $B_\epsilon(x) = \{y \in X \mid d(x, y) \leq \epsilon\}$ the *ball* of radius ϵ around x in (X, d) . The space (X, d) is *totally bounded* if for every $\epsilon > 0$ there exists a finite ϵ -*cover*, i.e. finitely many elements $x_1, \dots, x_n \in X$ such that $X = \bigcup_{i=1}^n B_\epsilon(x_i)$.

Recall that a metric space is compact iff it is complete and totally bounded.

We next introduce the Wasserstein and Kantorovich distances, which coincide according to Kantorovich-Rubinstein duality. To this end, we first need the notion of a coupling of two probability distributions, from which the original distributions are factored out as marginals.

Definition 3.2. Let π_1 and π_2 be discrete probability measures on A and B , respectively. We denote by $\text{Cpl}(\pi_1, \pi_2)$ the set of *couplings* of π_1 and π_2 , i.e. probability measures μ on $A \times B$ such that π_1 and π_2 are *marginals* of μ :

- for all $a \in A$, $\sum_{b \in B} \mu(a, b) = \pi_1(a)$;
- for all $b \in B$, $\sum_{a \in A} \mu(a, b) = \pi_2(b)$.

Definition 3.3 (Wasserstein and Kantorovich distances). Let (X, d) be a pseudometric space. We generally write

$$\mathcal{DX}$$

for the set of discrete probability distributions on X . We define two pseudometrics on \mathcal{DX} , the *Kantorovich distance* d^\uparrow and the *Wasserstein distance* d^\downarrow :

$$d^\uparrow(\pi_1, \pi_2) = \bigvee \{ |E_{\pi_1}(f) - E_{\pi_2}(f)| \mid f \in \text{Pred}(X, d) \}$$

$$d^\downarrow(\pi_1, \pi_2) = \bigwedge \{ E_\mu(d) \mid \mu \in \text{Cpl}(\pi_1, \pi_2) \}$$

where \bigwedge takes meets (and \bigvee suprema). We extend these distances without further mention to zero functions (like the functions r_a at blocking states a) by decreeing that the zero function has distance 1 from all probability distributions.

The notation d^\uparrow, d^\downarrow is meant as a mnemonic for the fact that these distances are obtained via suprema respectively via infima. If (X, d) is separable (contains a countable dense subset), these pseudometrics coincide, a fact known as the *Kantorovich-Rubinstein duality* (e.g. [Dudley, 2002]):

Lemma 3.4 (Kantorovich-Rubinstein duality). *Let (X, d) be a separable pseudometric space. Then for all $\pi_1, \pi_2 \in \mathcal{DX}$,*

$$d^\uparrow(\pi_1, \pi_2) = d^\downarrow(\pi_1, \pi_2).$$

The above notions of *lifting* a distance on X to a distance on distributions over X can be used to give fixed point equations for behavioural distances on models.

Definition 3.5 (Fixed point iteration à la Wasserstein/Kantorovich). Given a model \mathcal{I} , we define the chains $(d_n^K), (d_n^W)$ of *depth- n Kantorovich* and *Wasserstein distances*, respectively, via fixed point iteration:

$$d_0^W(a, b) = d_0^K(a, b) = 0$$

$$d_{n+1}^W(a, b) = \bigvee_{A \in \mathbf{Nc}} |A^{\mathcal{I}}(a) - A^{\mathcal{I}}(b)| \vee (d_n^W)^\downarrow(\pi_a, \pi_b)$$

$$d_{n+1}^K(a, b) = \bigvee_{A \in \mathbf{Nc}} |A^{\mathcal{I}}(a) - A^{\mathcal{I}}(b)| \vee (d_n^K)^\uparrow(\pi_a, \pi_b)$$

where \vee is binary join. We extend this to states a, b in different models \mathcal{I}, \mathcal{J} by taking the disjoint union of \mathcal{I}, \mathcal{J} .

In both cases, we start with the zero pseudometric, and in the next iteration lift the pseudometric d_n from the previous step via Wasserstein/Kantorovich. This lifted metric is then applied to the probability distributions π_a, π_b associated with a, b . In addition we take the maximum with the supremum over the distances for all atomic $A \in \mathbf{N}_C$.

We now introduce a key tool for our technical development, a new up-to- ϵ bisimulation game inspired by the definition of the Wasserstein distance.

Definition 3.6 (Bisimulation game). Given models \mathcal{I}, \mathcal{J} , $a_0 \in \Delta^{\mathcal{I}}, b_0 \in \Delta^{\mathcal{J}}$, and $\epsilon_0 \in [0, 1]$, the ϵ_0 -bisimulation game for a_0 and b_0 is played by Spoiler (S) and Duplicator (D), with rules as follows:

- *Configurations*: triples (a, b, ϵ) , with states $a \in \Delta^{\mathcal{I}}, b \in \Delta^{\mathcal{J}}$ and maximal allowed deviation $\epsilon \in [0, 1]$. The *initial configuration* is (a_0, b_0, ϵ_0) .
- *Moves*: In each round, D first picks a probability measure $\mu \in \text{Cpl}(\pi_a, \pi_b)$. Then, D distributes the deviation ϵ over all pairs (a', b') of successors, i.e. picks a function $\epsilon' : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{J}} \rightarrow [0, 1]$ such that $\mathbb{E}_\mu(\epsilon') \leq \epsilon$. Finally, S picks a pair (a', b') with $\mu(a', b') > 0$; the new configuration is then $(a', b', \epsilon'(a', b'))$.
- D wins if both states are blocking or $\epsilon = 1$.
- S wins if exactly one state is blocking and $\epsilon < 1$.
- *Winning condition*: $|A^{\mathcal{I}}(a) - A^{\mathcal{J}}(b)| \leq \epsilon$ for all $A \in \mathbf{N}_C$.

The game comes in two variants, the (unbounded) bisimulation game and the n -round bisimulation game, where $n \geq 0$. Player D wins if the winning condition holds *before* every round, otherwise S wins. More precisely, D wins the unbounded game if she can force an infinite play and the n -round game once n rounds have been played (the winning condition is not checked after the last round, so in particular, any 0-round game is an immediate win for D).

Remark 3.7. The above bisimulation game differs from bisimulation games in the literature (e.g. [Desharnais *et al.*, 2008]) in a number of salient features. A particularly striking aspect is that D 's moves are not similar to those of S , and moreover D in fact moves before S . Intuitively, D is required to commit beforehand to a strategy that she will use to respond to S 's next move. Note also that the precision ϵ changes as the game is being played, a complication forced by the arithmetic nature of models.

This leads to notions of game distance:

Definition 3.8. *depth- n game distance* d_n^G and (unbounded-depth) *game distance* d^G are defined as

$$d_n^G(a, b) = \bigwedge \{ \epsilon \mid D \text{ wins } G_n(a, b, \epsilon) \}$$

$$d^G(a, b) = \bigwedge \{ \epsilon \mid D \text{ wins } G(a, b, \epsilon) \}.$$

where $G(a, b, \epsilon)$ and $G_n(a, b, \epsilon)$ denote the the bisimulation game and the n -round bisimulation game on (a, b, ϵ) , respectively.

Finally we define the depth- n logical distance via $\mathcal{ALC}(\mathbf{P})$, restricting to concepts of rank at most n :

Definition 3.9. The *depth- n logical distance* $d_n^L(a, b)$ of states a, b in models \mathcal{I}, \mathcal{J} is defined as

$$d_n^L(a, b) = \bigvee \{ |C^{\mathcal{I}}(a) - C^{\mathcal{J}}(b)| \mid \text{rk}(C) \leq n \}.$$

The equivalence of the four bounded-depth behavioural distances introduced above will be shown in Theorem 4.3.

Behavioural distance forms the yardstick for our notion of bisimulation invariance; for definiteness:

Definition 3.10. A quantitative, i.e. $[0, 1]$ -valued, property Q of states, or a formula or concept defining such a property, is *bisimulation-invariant* if Q is non-expansive wrt. game distance, i.e. for states a, b in models \mathcal{I}, \mathcal{J} , respectively,

$$|Q(a) - Q(b)| \leq d^G(a, b).$$

Similarly, Q is *depth- n bisimulation invariant*, or *finite-depth bisimulation invariant* if mention of n is omitted, if Q is non-expansive wrt. d_n^G in the same sense.

It is easy to see that $\mathcal{ALC}(\mathbf{P})$ -concepts are bisimulation-invariant. More precisely, $\mathcal{ALC}(\mathbf{P})$ -concepts of rank at most n are depth- n bisimulation invariant (a stronger invariance since clearly $d_n^G \leq d^G$), as shown by routine induction. In contrast, many other properties of states are expressible in $\text{FO}(\mathbf{P})$ but not in $\mathcal{ALC}(\mathbf{P})$, as they fail to be bisimulation-invariant. Examples include $x\mathbf{P}[y : x = y]$ (probability of a self-transition) and $\exists z. x\mathbf{P}[y : y = z]$ (highest transition probability to a successor).

We are now ready to formally state our main theorem (a proof will be given in Section 6):

Theorem 3.11 (Modal characterization). *Every bisimulation-invariant $\text{FO}(\mathbf{P})$ -formula of rank at most n can be approximated (uniformly across all models) by $\mathcal{ALC}(\mathbf{P})$ -concepts of rank at most 3^n .*

(The exponential bound on the rank features also in the full statement of van Benthem's theorem.)

4 Modal Approximation at Finite Depth

We now establish the most important stepping stone on the way to the eventual proof of the modal characterization theorem: We show that every depth- n bisimulation-invariant property of states can be approximated by $\mathcal{ALC}(\mathbf{P})$ -concepts of rank at most n . We prove this simultaneously with coincidence of the various finite-depth behavioural pseudometrics defined in the previous section. To begin,

Lemma 4.1. *The game-based pseudometric d_n^G coincides with the Wasserstein pseudometric d_n^W ,*

We note next that the modality \mathbf{P} is non-expansive: We extend \mathbf{P} to act on $[0, 1]$ -valued functions $f : \Delta^{\mathcal{I}} \rightarrow [0, 1]$ by

$$(\mathbf{P}f)(a) = \mathbb{E}_{r_a}(f).$$

Lemma 4.2. *The map $f \mapsto \mathbf{P}f$ is non-expansive wrt. the supremum metric, that is $\|\mathbf{P}f - \mathbf{P}g\|_\infty \leq \|f - g\|_\infty$ for all $f, g : \Delta^{\mathcal{I}} \rightarrow [0, 1]$.*

Following our previous work [Wild *et al.*, 2018], we prove coincidence of the remaining pseudometrics in one big induction, along with total boundedness (needed later to apply a variant of the Arzelà-Ascoli theorem and the Kantorovich-Rubinstein duality) and modal approximability of depth- n bisimulation-invariant properties. We phrase the latter as density of the (semantics of) $\mathcal{ALC}(\mathbf{P})$ -concepts of rank at most n in the non-expansive function space (Definition 3.1):

Theorem 4.3. *Let \mathcal{I} be a model. Then for all $n \geq 0$,*

1. *we have $d_n^G = d_n^W = d_n^K = d_n^L =: d_n$ on \mathcal{I} ;*
2. *the pseudometric space $(\Delta^{\mathcal{I}}, d_n)$ is totally bounded;*
3. *$\mathcal{ALC}(\mathbf{P})_n$ is a dense subset of $\text{Pred}(\Delta^{\mathcal{I}}, d_n)$.*

Proof sketch. By simultaneous induction on n .

In the base case $n = 0$, all the behavioural distances are the zero pseudometric, so that total boundedness follows trivially and the density claim follows because non-expansive maps are just constants in $[0, 1]$ and the syntax of $\mathcal{ALC}(\mathbf{P})$ includes truth constants $q \in \mathbb{Q} \cap [0, 1]$.

For the inductive step, let \mathcal{I} be a model and $n > 0$, and assume as the inductive hypothesis that all claims in Theorem 4.3 hold for all $n' < n$. We begin with Item 1; $d_n^G = d_n^W$ is already proved (Lemma 4.1).

- $d_n^W = d_n^K$ follows by Kantorovich-Rubinstein duality (Lemma 3.4), since every totally bounded pseudometric space is separable.

- $d_n^K = d_n^L$: By Lemma 4.2 and the inductive hypothesis, $\mathbf{P}[\mathcal{ALC}(\mathbf{P})_{n-1}]$ is dense in $\mathbf{P}[\text{Pred}(\Delta^{\mathcal{I}}, d_{n-1})]$. Thus, the supremum in the definition of d_n^K does not change when it is taken only over the concepts in $\mathcal{ALC}(\mathbf{P})_{n-1}$ instead of all nonexpansive properties. The proof is finished by a simple induction over propositional combinations of concepts.

Item 2: By the inductive hypothesis, the space $(\Delta^{\mathcal{I}}, d_{n-1})$ is totally bounded. By the Arzelà-Ascoli theorem (in a version for totally bounded spaces and non-expansive maps, cf. [Wild et al., 2018]), it follows that $\text{Pred}(\Delta^{\mathcal{I}}, d_{n-1})$ is totally bounded wrt. the supremum pseudometric. This implies that depth- n distances can be approximated up to ϵ by examining differences at only finitely many, say m , concepts. As $([0, 1]^m, d_\infty)$ is totally bounded, $(\Delta^{\mathcal{I}}, d_n)$ is, too.

Item 3: By the Stone-Weierstraß theorem (again in a version for totally bounded spaces and non-expansive maps [Wild et al., 2018]) it suffices to give, for each $\epsilon > 0$, each non-expansive map $f \in \text{Pred}(\Delta^{\mathcal{I}}, d_n)$, and each pair of states $a, b \in \Delta^{\mathcal{I}}$ a concept $C \in \mathcal{ALC}(\mathbf{P})_n$ such that

$$\max(|f(a) - C^{\mathcal{I}}(a)|, |f(b) - C^{\mathcal{I}}(b)|) \leq \epsilon.$$

To construct such a C , we note that $|f(a) - f(b)| \leq d_n^L(a, b)$ (by non-expansiveness), so there exists some $D \in \mathcal{ALC}(\mathbf{P})_n$ such that $|D^{\mathcal{I}}(a) - D^{\mathcal{I}}(b)| \geq |f(a) - f(b)| - \epsilon$. From D , we can construct C using truncated subtraction \ominus . \square

This completes the proof of Theorem 4.3. Now that we can approximate depth- k bisimulation-invariant properties by $\mathcal{ALC}(\mathbf{P})$ -concepts of rank k on any fixed model, we need to make the approximation uniform across all models. We achieve this by means of a *final* model, i.e. one that realizes all behaviours. Formally:

Definition 4.4. A (probabilistic) bounded morphism between models \mathcal{I}, \mathcal{J} is a map $f : \Delta^{\mathcal{I}} \rightarrow \Delta^{\mathcal{J}}$ such that $A^{\mathcal{I}} = f^{-1}[A^{\mathcal{J}}]$ for each $A \in \mathbf{N}_C$ and $r_{f(a)}(B) = r_a(f^{-1}[B])$ for all $B \subseteq \Delta^{\mathcal{J}}$, $a \in \Delta^{\mathcal{I}}$ (implying that a is blocking iff $f(a)$ is blocking). A model \mathcal{F} is *final* if for every model \mathcal{I} , there exists a unique bounded morphism $\mathcal{I} \rightarrow \mathcal{F}$.

It follows from standard results in coalgebra [Barr, 1993] that a final model exists. Bounded morphisms preserve behaviour on-the-nose, that is:

Lemma 4.5. *Let $f : \mathcal{I} \rightarrow \mathcal{J}$ be a bounded morphism. Then, for any $a \in \Delta^{\mathcal{I}}$, $d^G(a, f(a)) = 0$.*

This entails the following lemma, which will enable us to use approximants on the final model as uniform approximants across all models:

Lemma 4.6. *Let \mathcal{F} be a final model, and let ϕ and ψ be bisimulation-invariant first-order properties. Then, for any model \mathcal{I} , $\|\phi - \psi\|_\infty^{\mathcal{I}} \leq \|\phi - \psi\|_\infty^{\mathcal{F}}$.*

5 Locality

The proof of the modal characterization theorem now further proceeds by first establishing that every bisimulation-invariant first-order formula ϕ is *local* in a sense to be made precise shortly, and subsequently that ϕ is in fact even finite-depth bisimulation invariant, for a depth that is exponential in the rank of ϕ . Locality refers to a probabilistic variant of Gaifman graphs [Gaifman, 1982]:

Definition 5.1. Let \mathcal{I} be a model.

- The *Gaifman graph* of \mathcal{I} is the undirected graph on the set $\Delta^{\mathcal{I}}$ of vertices that has an edge for every pair (a, b) with $r^{\mathcal{I}}(a, b) > 0$ or $r^{\mathcal{I}}(b, a) > 0$.

- The *Gaifman distance* $D : \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \rightarrow \mathbb{N} \cup \{\infty\}$ is graph distance in the Gaifman graph: For every $a, b \in \Delta^{\mathcal{I}}$, the distance $D(a, b)$ is the least number of edges on any path from a to b , if such a path exists, and ∞ otherwise.

- For $a \in \Delta^{\mathcal{I}}$ and $k \geq 0$, the *radius k neighbourhood* $U^k(a) = \{b \in \Delta^{\mathcal{I}} \mid D(a, b) \leq k\}$ of a consists of the states reachable from a in at most k steps.

- The *restriction* of \mathcal{I} to $U^k(a)$ is the model \mathcal{I}_a^k with set $U^k(a)$ of states, and

$$A^{\mathcal{I}_a^k}(b) = A^{\mathcal{I}}(b) \quad r^{\mathcal{I}_a^k}(b, c) = \begin{cases} r^{\mathcal{I}}(b, c) & \text{if } D(a, b) < k \\ 0 & \text{if } D(a, b) = k \end{cases}$$

for $A \in \mathbf{N}_C$ and $b, c \in U^k(a)$.

The restriction to $U^k(a)$ thus makes all states at distance k blocking. Restricted models have the expected relationship with games of bounded depth:

Lemma 5.2. *Let a be a state in a model \mathcal{I} . Then D wins the k -round 0-bisimulation game for \mathcal{I} , a and \mathcal{I}_a^k, a .*

Locality of a formula now means that its truth values only depend on the neighbourhood of the state in question:

Definition 5.3. A formula $\phi(x)$ is *k -local* for a radius k if for every model \mathcal{I} and every $a \in \Delta^{\mathcal{I}}$, $\phi^{\mathcal{I}}(a) = \phi^{\mathcal{I}_a^k}(a)$.

As $\mathcal{ALC}(\mathbf{P})$ -concepts are bisimulation-invariant, Lemma 5.2 implies

Lemma 5.4. *Every $\mathcal{ALC}(\mathbf{P})$ -concept of rank at most k is k -local.*

To prove locality of bisimulation-invariant FO(\mathbf{P})-formulae, we require a model-theoretic tool, an adaptation of Ehrenfeucht-Fraïssé equivalence to the probabilistic setting:

Definition 5.5. Let \mathcal{I}, \mathcal{J} be models, and let \bar{a}_0 and \bar{b}_0 be vectors of equal length over $\Delta^{\mathcal{I}}$ and $\Delta^{\mathcal{J}}$, respectively. The Ehrenfeucht-Fraïssé game for \mathcal{I}, \bar{a}_0 and \mathcal{J}, \bar{b}_0 , played by Spoiler (S) and Duplicator (D), is given as follows.

- *Configurations:* pairs (\bar{a}, \bar{b}) of vectors \bar{a} over $\Delta^{\mathcal{I}}$ and \bar{b} over $\Delta^{\mathcal{J}}$; the initial configuration is (\bar{a}_0, \bar{b}_0) .
- *Moves:* Each round can be played in one of two ways, chosen by S :
 - *Standard round:* S selects a state in one model, say $a \in \Delta^{\mathcal{I}}$, and D then has to select a state in the other model, say $b \in \Delta^{\mathcal{J}}$, reaching the configuration $(\bar{a}a, \bar{b}b)$.
 - *Probabilistic round:* S selects an index i and a fuzzy subset in one model, say $\phi_A: \Delta^{\mathcal{I}} \rightarrow [0, 1]$. D then has to select a fuzzy subset in the other model, say $\phi_B: \Delta^{\mathcal{J}} \rightarrow [0, 1]$, such that $E_{r_{a_i}}(\phi_A) = E_{r_{b_i}}(\phi_B)$. Then, S selects an element on one side, say $a \in \Delta^{\mathcal{I}}$, such that $r_{a_i}(a) > 0$, and D subsequently selects an element on the other side, say $b \in \Delta^{\mathcal{J}}$, such that $\phi_A(a) = \phi_B(b)$ and $r_{b_i}(b) > 0$, reaching the configuration $(\bar{a}a, \bar{b}b)$.
- *Winning conditions:* Any player who cannot move loses. S wins if a configuration is reached (including the initial configuration) that fails to be a partial isomorphism. Here, a configuration (\bar{a}, \bar{b}) is a *partial isomorphism* if
 - $a_i = a_j \iff b_i = b_j$
 - $A^{\mathcal{I}}(a_i) = A^{\mathcal{J}}(b_i)$ for all i and all $A \in \mathcal{N}_C$
 - $r^{\mathcal{I}}(a_i, a_j) = r^{\mathcal{J}}(b_i, b_j)$ for all i, j .

Player D wins if she reaches the n -th round (maintaining configurations that are not winning for S).

For our purposes, we need only soundness of Ehrenfeucht-Fraïssé equivalence:

Lemma 5.6 (Ehrenfeucht-Fraïssé invariance). *Let \mathcal{I}, \mathcal{J} be models, and let \bar{a}_0, \bar{b}_0 be vectors of length m over $\Delta^{\mathcal{I}}$ and $\Delta^{\mathcal{J}}$, respectively, such that D wins the n -round Ehrenfeucht-Fraïssé game on \bar{a}_0, \bar{b}_0 . Then for every FO(\mathcal{P})-formula ϕ with $\text{qr}(\phi) \leq n$ and free variables at most x_1, \dots, x_m ,*

$$\phi(\bar{a}_0) = \phi(\bar{b}_0).$$

Since embeddings into disjoint unions of models are bounded morphisms, the following is immediate from Lemma 4.5:

Lemma 5.7. *Every bisimulation-invariant formula is also invariant under disjoint union.*

We are now in a position to prove our desired locality result:

Lemma 5.8 (Locality). *Let $\phi(x)$ be a bisimulation-invariant FO(\mathcal{P})-formula of rank n with one free variable x . Then ϕ is k -local for $k = 3^n$.*

Proof sketch. Let a be a state in a model \mathcal{I} . We need to show $\phi^{\mathcal{I}}(a) = \phi^{\mathcal{I}_a^k}(a)$. Construct models \mathcal{J}, \mathcal{K} that extend \mathcal{I} and \mathcal{I}_a^k , respectively, by adding n disjoint copies of both \mathcal{I} and \mathcal{I}_a^k . We finish the proof by showing that

$$\phi^{\mathcal{I}}(a) = \phi^{\mathcal{J}}(a) = \phi^{\mathcal{K}}(a) = \phi^{\mathcal{I}_a^k}(a).$$

The first and third equality follow by bisimulation invariance of ϕ (Lemma 5.7), and the second using Lemma 5.6, by giving a winning invariant for D in the n -round Ehrenfeucht-Fraïssé game for \mathcal{J}, a and \mathcal{K}, a . \square

6 Proof of the Main Result

Having established locality of bisimulation-invariant first-order formulae and modal approximability of finite-depth bisimulation-invariant properties, we now discharge the last remaining steps in our programme: We show by means of an unravelling construction that bisimulation-invariant first-order formulae are already finite-depth bisimulation-invariant, and then conclude the proof of our main result, the modal characterization theorem.

Definition 6.1. Let \mathcal{I} be a model. The *unravelling* \mathcal{I}^* of \mathcal{I} is a model with non-empty finite sequences $\bar{a} \in (\Delta^{\mathcal{I}})^+$ as states, where atomic concepts and roles are interpreted by

$$A^{\mathcal{I}^*}(\bar{a}) = A^{\mathcal{I}}(\text{last}(\bar{a})) \quad r^{\mathcal{I}^*}(\bar{a}, \bar{a}a) = r^{\mathcal{I}}(\text{last}(\bar{a}), a),$$

for $\bar{a} \in (\Delta^{\mathcal{I}})^+$ and $a \in \Delta^{\mathcal{I}}$, where *last* takes last elements.

As usual, models are bisimilar to their unravellings:

Lemma 6.2. *For any model \mathcal{I} and $a \in \Delta^{\mathcal{I}}$, D has a winning strategy in the 0-bisimulation game for \mathcal{I}, a and \mathcal{I}^*, a .*

We next show that locality and bisimulation invariance imply finite-depth bisimulation invariance:

Lemma 6.3. *Let ϕ be bisimulation invariant and k -local. Then ϕ is depth- k bisimulation invariant.*

Proof sketch. By unravelling (Lemma 6.2) and locality (Lemma 5.2), we need only consider depth- k tree models. On such models, winning strategies in k -round bisimulation games automatically win also the unrestricted game. \square

This allows us to wrap up the proof of our main result:

Proof of Theorem 3.11. Let ϕ be a probabilistic first-order formula of rank n . By Lemma 5.8 and Lemma 6.3, ϕ is depth- k bisimulation-invariant for $k = 3^n$. By Theorem 4.3, for every $\epsilon > 0$, there exists an $\mathcal{ALC}(\mathcal{P})$ concept C_ϵ of rank at most k such that $\|\phi^{\mathcal{F}} - C_\epsilon^{\mathcal{F}}\|_\infty \leq \epsilon$ on the final model \mathcal{F} . By Lemma 4.6, this approximation works over all models. \square

7 Conclusions

We have established a modal characterization result for a probabilistic fuzzy DL $\mathcal{ALC}(\mathcal{P})$, stating that every formula of quantitative probabilistic FOL that is *bisimulation-invariant*, i.e. non-expansive wrt. a natural notion of behavioural distance, can be approximated by $\mathcal{ALC}(\mathcal{P})$ -concepts of bounded modal rank, the bound being exponential in the rank of the original formula. As discussed in the introduction, the bound on the modal rank is the crucial feature making this result into a van-Benthem (rather than Hennessy-Milner) type theorem.

It remains open whether our main result can be sharpened to make do without approximation. (Similar open problems persist for the case of fuzzy modal logic [Wild *et al.*, 2018] and two-valued probabilistic modal logic [Schroder *et al.*, 2017].) Further directions for future research include a treatment of Łukasiewicz semantics of the propositional connectives (for which non-expansiveness in fact fails). Moreover, the version of our main result that restricts the semantics to finite models, in analogy to Rosen’s finite-model version of van Benthem’s theorem [Rosen, 1997], remains open.

References

- [Baldan *et al.*, 2014] Paolo Baldan, Filippo Bonchi, Henning Kerstan, and Barbara König. Behavioral metrics via functor lifting. In *Found. Software Tech. Theor. Comput. Sci., FSTTCS 2014*, volume 29 of *LIPICs*, pages 403–415, 2014.
- [Barr, 1993] Michael Barr. Terminal coalgebras in well-founded set theory. *Theor. Comp. Sci.*, 114:299–315, 1993.
- [Burgess, 1969] John Burgess. Probability logic. *J. Symb. Log.*, 34:264–274, 1969.
- [Desharnais *et al.*, 2004] Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labelled Markov processes. *Theor. Comput. Sci.*, 318:323–354, 2004.
- [Desharnais *et al.*, 2008] Josée Desharnais, François Laviolette, and Mathieu Tracol. Approximate analysis of probabilistic processes: Logic, simulation and games. In *Quantitative Evaluation of Systems, QEST 2008*, pages 264–273. IEEE, 2008.
- [Dudley, 2002] Richard Dudley. *Real Analysis and Probability*. Cambridge University Press, 2nd edition, 2002.
- [Figueira *et al.*, 2015] Diego Figueira, Santiago Figueira, and Carlos Areces. Model theory of XPath on data trees. Part I: Bisimulation and characterization. *J. Artif. Intell. Res. (JAIR)*, 53:271–314, 2015.
- [Flaminio and Godo, 2007] Tommaso Flaminio and Lluís Godo. A logic for reasoning about the probability of fuzzy events. *Fuzzy Sets and Systems*, 158(6):625–638, 2007.
- [Gaifman, 1982] Haim Gaifman. On local and non-local properties. In *Logic Colloq. 1981*, pages 105–135, 1982.
- [Giacalone *et al.*, 1990] Alessandro Giacalone, Chi Jou, and Scott Smolka. Algebraic reasoning for probabilistic concurrent systems. In *Programming Concepts and Methods*, pages 443–458. North-Holland, 1990.
- [Hájek, 2007] Petr Hájek. Complexity of fuzzy probability logics II. *Fuzzy Sets and Systems*, 158:2605–2611, 2007.
- [Halpern and Rabin, 1987] Joseph Halpern and Michael Rabin. A logic to reason about likelihood. *Artif. Intell.*, 32:379–405, 1987.
- [Halpern, 1990] Joseph Halpern. An analysis of first-order logics of probability. *Artif. Intell.*, 46:311–350, 1990.
- [Hamblin, 1959] Charles Hamblin. The modal ‘probably’. *Mind*, 68:234–240, 1959.
- [Heifetz and Mongin, 2001] Aviad Heifetz and Philippe Mongin. Probabilistic logic for type spaces. *Games Econ. Behav.*, 35:31–53, 2001.
- [Hennessy and Milner, 1985] Matthew Hennessy and Robin Milner. Algebraic laws for non-determinism and concurrency. *J. ACM*, 32:137–161, 1985.
- [Herzig, 2003] Andreas Herzig. Modal probability, belief, and actions. *Fund. Inf.*, 57:323–344, 2003.
- [Huth and Kwiatkowska, 1997] Michael Huth and Marta Kwiatkowska. Quantitative analysis and model checking. In *Logic in Computer Science, LICS 1997*, pages 111–122. IEEE, 1997.
- [Janin and Walukiewicz, 1995] David Janin and Igor Walukiewicz. Automata for the modal μ -calculus and related results. In *Mathematical Foundations of Computer Science, MFCS 1995*, volume 969 of *LNCS*, pages 552–562. Springer, 1995.
- [Kozen, 1985] Dexter Kozen. A probabilistic PDL. *J. Comput. Sys. Sci.*, 30(2):162–178, 1985.
- [Larsen and Skou, 1991] Kim Larsen and Arne Skou. Bisimulation through probabilistic testing. *Inf. Comput.*, 94:1–28, 1991.
- [Litak *et al.*, 2018] Tadeusz Litak, Dirk Pattinson, Katsuhiko Sano, and Lutz Schröder. Model theory and proof theory of coalgebraic predicate logic. *Log. Meth. Comput. Sci.*, 14(1), 2018.
- [Lukasiewicz and Straccia, 2008] Thomas Lukasiewicz and Umberto Straccia. Managing uncertainty and vagueness in description logics for the semantic web. *J. Web Sem.*, 6:291–308, 2008.
- [Makowsky and Ziegler, 1980] Johann Makowsky and Martin Ziegler. Topological model theory with an interior operator: consistency properties and back - and forth arguments. *Arch. math. Logik*, 20:27–54, 1980.
- [Otto, 2004] Martin Otto. Elementary proof of the van Benthem-Rosen characterisation theorem. Technical Report 2342, TU Darmstadt, 2004.
- [Rosen, 1997] Eric Rosen. Modal logic over finite structures. *J. Logic Lang. Inform.*, 6(4):427–439, 1997.
- [Schröder and Pattinson, 2011] Lutz Schröder and Dirk Pattinson. Description logics and fuzzy probability. In *IJCAI 2011*, pages 1075–1081, 2011.
- [Schröder *et al.*, 2017] Lutz Schröder, Dirk Pattinson, and Tadeusz Litak. A van Benthem/Rosen theorem for coalgebraic predicate logic. *J. Log. Comput.*, 27:749–773, 2017.
- [Sturm and Wolter, 2001] Holger Sturm and Frank Wolter. First-order expressivity for S5-models: Modal vs. two-sorted languages. *J. Philos. Logic*, 30:571–591, 2001.
- [van Benthem, 1976] Johan van Benthem. *Modal Correspondence Theory*. PhD thesis, Univ. of Amsterdam, 1976.
- [van Breugel and Worrell, 2005] Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331:115–142, 2005.
- [Wild and Schröder, 2017] Paul Wild and Lutz Schröder. A characterization theorem for a modal description logic. In *IJCAI 2017*, pages 1304–1310, 2017.
- [Wild *et al.*, 2018] Paul Wild, Lutz Schröder, Dirk Pattinson, and Barbara König. A van Benthem theorem for fuzzy modal logic. In *Logic in Computer Science, LICS 2018*, pages 909–918. ACM, 2018.
- [Zadeh, 1968] Lotfi Zadeh. Probability measures of fuzzy events. *J. Math. Anal. Appl.*, 23:421–427, 1968.