

On the Learnability of Possibilistic Theories

Cosimo Persia and Ana Ozaki*

University of Bergen

Abstract

We investigate learnability of possibilistic theories from entailments in light of Angluin’s exact learning model. We consider cases in which only membership, only equivalence, and both kinds of queries can be posed by the learner. We then show that, for a large class of problems, polynomial time learnability results for classical logic can be transferred to the respective possibilistic extension. In particular, it follows from our results that the possibilistic extension of propositional Horn theories is exactly learnable in polynomial time. As polynomial time learnability in the exact model is transferable to the classical probably approximately correct model extended with membership queries, our work also establishes such results in this model.

1 Introduction

Uncertainty is found in many phases of learning, such as model selection and processing noisy, imperfect, incomplete or limited data. In most cases, knowledge-based systems are constrained to live under conditions of ignorance. There are different approaches to deal with uncertainty [Parsons and Hunter, 1998]. A well-studied formalism for dealing with it is *possibilistic logic* [Didier Dubois and Prade, 1994; Lang, 2000]. It admits a graded notion of possibility and makes a clear distinction between the concepts of truth and belief [Dubois and Prade, 2001]. Uncertainty of formulas in possibilistic logic is not subject to the complement rule as in probability theory [Agarwal and Nayal, 2015; Dubois and Prade, 1993]. Indeed, complementary formulas may be considered fully possible, meaning complete ignorance about their truth value.

Example 1. Consider a doctor who has to diagnose a patient that suffers from extreme fatigue. A doctor can consider blood-related conditions: iron deficiency, iron overload, and vitamin B12 deficiency. Within possibility theory, one can model cases of complete uncertainty. Both iron deficiency and iron overload, which are two mutually exclusive conditions, can be considered fully possible. Consider that vitamin B12 deficiency is considered to be less possible, e.g. associated

with the value $1/3$, based on some information provided by the patient. In probability theory, complete ignorance of the first two conditions would make us assign probability $1/3$ to every condition (Laplace criterion). Thus, it would not model the knowledge about vitamin B12 deficiency and the ignorance about iron deficiency and iron overload. ◀

Although possibilistic logic has been extensively studied [Dubois and Prade, 2015], there are not many works that investigate learnability of possibilistic theories. In this work, we partially cover this gap by studying whether possibilistic theories are learnable in Angluin’s exact learning model [Angluin, 1988]. In this model, a learner interacts with a teacher to exactly identify an abstract target concept. One can see the doctor, in Example 1, as a learner who inquires the patient (playing the role of a teacher) in order to identify a disease.

The most studied communication protocol in this model contains queries of two kinds, called *membership* and *equivalence* queries. Membership queries allow the learner to know whether a certain statement holds. Equivalence queries allow the learner to check whether a hypothesis (e.g. a diagnose) is correct and, if not, to fix it using a counterexample. In our toy scenario, the patient may not be able to provide a counterexample but new symptoms or reactions can reveal that the hypothesis is not correct. To the best of our knowledge, this is the first work where learnability of possibilistic theories is investigated in Angluin’s model. We consider cases in which only membership, only equivalence, and both kinds of queries can be posed by the learner. We also study whether known polynomial time exact learning results for classical logic can be transferred to possibilistic settings.

Our main result is that, for a large class of problems, polynomial time learnability (with both types of queries) can be transferred from classical logic to the respective possibilistic extension (Theorem 17). If only membership queries are allowed (and the maximal precision of valuations in the target is fixed) then polynomial time learnability of a classical logic can also be transferred to the possibilistic extension. We leave open the case in which only equivalence queries can be asked. With our main result, we establish, e.g., that the possibilistic extension of propositional Horn [Angluin *et al.*, 1992; Frazier and Pitt, 1993; Hermo and Ozaki, 2020] and fragments of first-order Horn [Arimura, 1997; Reddy and Tadepalli, 1998; Konev *et al.*, 2018] are exactly learnable in polynomial time. As polynomial time learnability in the exact model is trans-

*Contact Author: Ana.Ozaki@uib.no

ferable to the probably approximately correct (PAC) [Valiant, 1984] model extended with membership queries, our work also establishes such results in this model.

Related Work. Among the works that combine learning and possibilistic logic, we can find results on learning possibilistic logic theories from default rules within the PAC learning model [Kuzelka *et al.*, 2016]. Possibilistic logic has been used to reason with default rules [Benferhat *et al.*, 1992] to select the most plausible rule and in inductive logic programming to handle exceptions [Serrurier and Prade, 2007]. In statistical relational learning, possibilistic logic has been used as a formal encoding of statistical regularities found in relational data [Kuzelka *et al.*, 2017]. Possibilistic formulas can encode Markov logic networks [Kuzelka *et al.*, 2015]. Formal concept analysis has been applied to generate attribute implications with a degree of certainty [Djouadi *et al.*, 2010]. We also point out an extension of version space learning that deals with examples associated with possibility degrees [Prade and Serrurier, 2008].

In Section 2, we present basic definitions. In Section 3, we investigate whether possibilistic logic theories can be learned and, in Section 4, we show transferability of polynomial time learnability results.

2 Basics

In the following, we provide relevant notions of possibilistic logic and learning theory used in the paper.

2.1 Possibilistic Theories

Let L be a propositional or a first-order (FO) language (restricted to well-formed formulas without free variables) with the semantics of classical FO logic. We say that $\varphi \in L$ is *satisfiable* if there is an interpretation \mathcal{I} such that φ is satisfied in \mathcal{I} . Moreover, φ is *falsifiable* if its negation $\neg\varphi$ is satisfiable. An *FO knowledge base* (FO KB) is a finite set of FO formulas. An FO KB is *non-trivial* if it is satisfiable and falsifiable. The *possibilistic extension* of an FO language L is defined as follows. A *possibilistic formula* is a pair (φ, α) , where $\varphi \in L$ and α is a real number (with finite precision) in the interval $(0, 1]$, called the *valuation* of φ . A *possibilistic KB* (or a possibilistic theory) is a finite set \mathcal{K} of possibilistic formulas. Given a set Ω of interpretations for L , a *possibility distribution* π is a function from Ω to the interval $[0, 1]$. The *possibility* and *necessity measures*, Π and N , are functions (induced by π) from L to $[0, 1]$, defined respectively as

$$\Pi(\varphi) = \sup\{\pi(\mathcal{I}) \mid \mathcal{I} \in \Omega, \mathcal{I} \models \varphi\}$$

$$N(\varphi) = 1 - \Pi(\neg\varphi) = \inf\{1 - \pi(\mathcal{I}) \mid \mathcal{I} \in \Omega, \mathcal{I} \models \neg\varphi\}.$$

A possibility distribution π *satisfies* a possibilistic formula (φ, α) , written $\pi \models (\varphi, \alpha)$, if $N(\varphi) \geq \alpha$, and it satisfies a possibilistic KB $\mathcal{K} = \{(\varphi_i, \alpha_i) \mid 0 \leq i < n\}$ if it satisfies each $(\varphi_i, \alpha_i) \in \mathcal{K}$. We have that (φ, α) is *entailed* by \mathcal{K} , written $\mathcal{K} \models (\varphi, \alpha)$, if all possibility distributions that satisfy \mathcal{K} also satisfy (φ, α) . Given \mathcal{K} as above and $\mathcal{I} \in \Omega$, we define the possibility distribution $\pi_{\mathcal{K}}$ as follows: $\pi_{\mathcal{K}}(\mathcal{I}) = 1$, if $\mathcal{I} \models \varphi_i$, for every $(\varphi_i, \alpha_i) \in \mathcal{K}$; otherwise, $\pi_{\mathcal{K}}(\mathcal{I}) = \min\{1 - \alpha_i \mid \mathcal{I} \models \neg\varphi_i, 0 \leq i < n\}$.

The *FO projection* of \mathcal{K} is the set $\mathcal{K}^* = \{\varphi_i \mid (\varphi_i, \alpha_i) \in \mathcal{K}\}$. The α -*cut* and the $\bar{\alpha}$ -*cut* of \mathcal{K} , with $\alpha \in (0, 1]$, are defined respectively as $\mathcal{K}_\alpha = \{(\varphi, \beta) \in \mathcal{K} \mid \beta \geq \alpha\}$ and $\mathcal{K}_{\bar{\alpha}} = \{(\varphi, \beta) \in \mathcal{K} \mid \beta > \alpha\}$. The set of all valuations occurring in \mathcal{K} is $\mathcal{K}^v = \{\alpha \mid (\varphi, \alpha) \in \mathcal{K}\}$. Moreover, $\text{val}(\varphi, \mathcal{K}) = \sup\{\alpha \mid \mathcal{K} \models (\varphi, \alpha)\}$ is the least upper bound of the valuations of formulas entailed by \mathcal{K} . Finally, the *inconsistency degree* of \mathcal{K} is defined as $\text{inc}(\mathcal{K}) = \sup\{\alpha \mid \mathcal{K} \models (\perp, \alpha)\}$.

Lemma 2. [Didier Dubois and Prade, 1994] *Let \mathcal{K} be a possibilistic KB. For every possibilistic formula (ϕ, α) ,*

1. $\mathcal{K} \models (\phi, \alpha)$ iff $\mathcal{K}_\alpha^* \models \phi$;
2. $\mathcal{K} \models (\phi, \alpha)$ iff $\alpha \leq \text{val}(\phi, \mathcal{K})$; and
3. $\mathcal{K} \models (\phi, \alpha)$ implies $\text{val}(\phi, \mathcal{K}) \in \mathcal{K}^v \cup \{1\}$.

Proof. Point 1 is a consequence of Propositions 3.5.2, 3.5.5, and 3.5.6, and Point 2 is Property 1 at page 453 in [Didier Dubois and Prade, 1994]. We argue about Point 3. By definition of $\pi_{\mathcal{K}}$, for all $\mathcal{I} \in \Omega$, $\pi_{\mathcal{K}}(\mathcal{I})$ is either 1 or $1 - \beta$ for some $\beta \in \mathcal{K}^v$. Let $N_{\mathcal{K}}$ be the necessity measure induced by $\pi_{\mathcal{K}}$. By definition of $N_{\mathcal{K}}$, $N_{\mathcal{K}}(\phi) = \inf\{1 - \pi_{\mathcal{K}}(\mathcal{I}) \mid \mathcal{I} \in \Omega, \mathcal{I} \models \neg\phi\}$. Then, $N_{\mathcal{K}}(\phi) \in \mathcal{K}^v \cup \{0, 1\}$ (recall that $\inf\{\}$ is 1, which is the case for tautologies). By the semantics of possibilistic logic, $N_{\mathcal{K}}(\phi) = \text{val}(\phi, \mathcal{K})$ [Didier Dubois and Prade, 1994, Corollary 3.2.3]. As (ϕ, α) is a possibilistic formula, $\alpha > 0$. So, by Point 2, $N_{\mathcal{K}}(\phi) = \text{val}(\phi, \mathcal{K}) \in \mathcal{K}^v \cup \{1\}$. \square

We denote by $=_p$ the operator that checks if two numbers are equal up to precision p . For example $0.124 =_2 0.12345$ but $0.124 \neq_3 0.12345$. Assume $\alpha \in (0, 1]$ has finite precision. We write $\text{prec}(\alpha)$ for the precision of α and $\text{prec}(t)$ for $\sup\{\text{prec}(\alpha) \mid (\phi, \alpha) \in t\}$. Given an interval I , we write I_p for the set containing all $\alpha \in I$ with $\text{prec}(\alpha) = p$.

Example 3. One can express (1) mutual exclusion of iron deficiency and iron overload and (2) lower necessity of iron overload to be the cause of fatigue than iron deficiency with the possibilistic KB $\{(\forall x(\text{IronDef}(x) \rightarrow \neg\text{IronOver}(x)), 1), (\forall x(\text{IronDef}(x) \rightarrow \text{Fatigue}(x)), 0.9), (\forall x(\text{IronOver}(x) \rightarrow \text{Fatigue}(x)), 0.8)\}$. \triangleleft

2.2 Learnability

In learning theory, examples are pieces of information that characterise an abstract target the learner wants to learn. We consider the problem of learning targets represented in decidable fragments of FO logic or in their possibilistic extensions. Examples in our case are formulas expressed in the chosen logic (in this context called ‘entailments’).

A *learning framework* \mathfrak{F} is a pair $(\mathcal{E}, \mathcal{L})$; where \mathcal{E} is a non-empty and countable set of *examples*, and \mathcal{L} is a non-empty and countable set of *concept representations* (also called *hypothesis space*). Each element l of \mathcal{L} is assumed to be represented using a finite set of symbols Σ_l (the *signature* of l). In all learning frameworks considered in this work, \mathcal{E} is a set of formulas and \mathcal{L} is a set of KBs (in a chosen language). We say that $e \in \mathcal{E}$ is a *positive example* for $l \in \mathcal{L}$ if $l \models e$ and a *negative example* for l if $l \not\models e$. Given a learning framework $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$, we are interested in the exact identification of a *target* $t \in \mathcal{L}$, by posing queries to oracles. Let $\text{MQ}_{\mathfrak{F}, t}$ be the oracle that takes as input some $e \in \mathcal{E}$ and returns ‘yes’ if

$t \models e$ and ‘no’ otherwise. A *membership query* is a call to the oracle $\text{MQ}_{\mathfrak{F},t}$. Given $t, h \in \mathcal{L}$, a *counterexample* for t and h is an example $e \in \mathcal{E}$ s.t. $t \models e$ and $h \not\models e$ (or vice-versa, $h \models e$ and $t \not\models e$). For every $t \in \mathcal{L}$, we denote by $\text{EQ}_{\mathfrak{F},t}$ an oracle that takes as input a *hypothesis* $h \in \mathcal{L}$ and returns ‘yes’ if $h \equiv t$ and a counterexample otherwise. There is no assumption regarding which counterexample is chosen by the oracle. An *equivalence query* is a call to $\text{EQ}_{\mathfrak{F},t}$.

Example 4. A blood test to check for vitamin B12 deficiency on patient 42 can be modelled with a call to $\text{MQ}_{\mathfrak{F},t}$ with $(\text{B12Def}(\text{patient.42}), \alpha)$ for some $\alpha \in (0, 1]$ as input (depending on the result and accuracy of the test). \triangleleft

A *learner* for $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ is a deterministic algorithm that, for a fixed but arbitrary $t \in \mathcal{L}$, takes Σ_t as input, is allowed to pose queries to $\text{MQ}_{\mathfrak{F},t}$ and $\text{EQ}_{\mathfrak{F},t}$ (without knowing the target t), and that eventually halts and outputs some $h \in \mathcal{L}$ with $h \equiv t$. This notion of an algorithm with access to oracles can be formalised using *learning systems* [Watanabe, 1990], where posing a query to an oracle means writing down the query in an (additional) communication tape, entering in a query state, and waiting. The oracle then writes the answer in the communication tape, enters in an answer state, and stops. After that, the learner resumes its execution and can now read the answer in the communication tape.

We say that \mathfrak{F} is (exactly) *learnable* if there is a learner for \mathfrak{F} and that \mathfrak{F} is *polynomial time learnable* if it is learnable by a learner A such that at every step (the time used by an oracle to write an answer is *not* taken into account) of computation the time used by A up to that step is bounded by a polynomial $p(|t|, |e|)$, where $t \in \mathcal{L}$ is the target and $e \in \mathcal{E}$ is the largest counterexample seen so far. We denote by $\text{PTIME}_{\mathfrak{L}}$ the class of learning frameworks which are polynomial time learnable and the complexity of the entailment problem is in PTIME^1 . We also consider cases in which the learner can only pose one type of query (only membership or only equivalence queries). Whenever this is the case we write this explicitly.

Let $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ be a learning framework where \mathcal{E} is a set of FO formulas and \mathcal{L} is a set of FO KBs. We call such \mathfrak{F} an *FO learning framework*. We say that \mathfrak{F} is *non-trivial* if \mathcal{L} contains a non-trivial FO KB; and that it is *safe* if $l \in \mathcal{L}$ implies that $l' \in \mathcal{L}$, for all $l' \subseteq l$. A *possibilistic extension* l_π of an FO KB l is a possibilistic KB obtained by adding a possibilistic valuation α to every formula $\varphi \in l$. The possibilistic extension \mathfrak{F}_π of \mathfrak{F} is the pair $(\mathcal{E}_\pi, \mathcal{L}_\pi)$ where \mathcal{L}_π is the set of all possibilistic extensions of each $l \in \mathcal{L}$, and \mathcal{E}_π is the set of all possibilistic formulas entailed by an element of \mathcal{L}_π .

We write \mathbb{N}^+ for the set of positive natural numbers. Given $p \in \mathbb{N}^+$, we denote by $\mathfrak{F}_\pi^p = (\mathcal{E}_\pi, \mathcal{L}_\pi^p)$ the result of removing from \mathcal{L}_π in \mathfrak{F}_π every $l \in \mathcal{L}_\pi$ that does not satisfy $\text{prec}(l) = p$.

Remark 1. Let $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ be an FO learning framework and let $t \in \mathcal{L}$ be the target. If a learner A has access to $\text{MQ}_{\mathfrak{F},t}$ then we can assume w.l.o.g. that all counterexamples returned

¹In general, non-trivial algorithms need to perform entailment checks to combine the information of the examples. So polynomial time learning algorithms are normally for logics in which the entailment problem is tractable. This is the case e.g. for the Horn results mentioned in the Introduction.

by $\text{EQ}_{\mathfrak{F},t}$ are positive: the learner can check whether each $\phi \in h$ is entailed by t . The same holds for \mathfrak{F}_π .

3 Learnability Results

We start by studying the problem of whether there is a learner for a learning framework such that it always terminates with a hypothesis equivalent to the target. The main difficulty in learning with only membership queries (even for plain FO settings) is that the learner would ‘not know’ whether it has found a formula equivalent to a (non-trivial) target.

Example 5. Let $\Phi_n := \exists x_1 \dots \exists x_n. \bigwedge_{0 \leq i < n} r(x_i, x_{i+1})$. A learner may ask membership queries of the form $\exists x_0 \Phi_n$ for an arbitrarily large n without being able to distinguish whether the target theory is $\exists x_0 \Phi_n$ or $\forall x_0 (\Phi_n \rightarrow \Phi_{n+1})$ (knowing the signature of the target theory does not help the learner). \triangleleft

For possibilistic theories, another difficulty arises even for the propositional case. As the precision of a formula can be arbitrarily high, the learner may not know when to stop (e.g., is the target $(p, 0.1)$? or $(p, 0.11)$?). Theorem 6 states that, except for trivial cases, learnability cannot be guaranteed.

Theorem 6. Let \mathfrak{F} be a non-trivial FO learning framework. \mathfrak{F}_π is not (exactly) learnable with only membership queries.

Sketch. The existence of a learner A for the possibilistic extension $\mathfrak{F}_\pi = (\mathcal{E}_\pi, \mathcal{L}_\pi)$ of a non-trivial learning framework \mathfrak{F} would imply the existence of a procedure that terminates in n steps. A would not distinguish between the elements of \mathcal{L}_π with precision higher than n . \square

If the precision of the target is known or fixed, learnability of an FO learning framework can be transferred to its possibilistic extension. We state this in Theorem 8. To show this theorem, we use the following technical result.

Lemma 7. Let t be a possibilistic KB. Let I be a set of valuations such that $t^v \subseteq I$. If for each $\alpha \in I$ there is some FO KB k_α^* such that $k_\alpha^* \equiv t_\alpha^*$ then $t \equiv \{(\phi, \alpha) \mid \phi \in k_\alpha^*, \alpha \in I\}$.

Proof. Let $h = \{(\phi, \alpha) \mid \phi \in k_\alpha^*, \alpha \in I\}$. Assume $h \not\models (\phi, \gamma)$. If $\gamma = 1$ and $\gamma \notin I$ then ϕ is a tautology. In this case, for all $\beta \in (0, 1]$, $t \models (\phi, \beta)$. Suppose this is not the case. By Points 2 and 3 of Lemma 2, $\gamma \leq \alpha$, $\alpha = \text{val}(\phi, h) \in h^v \cup \{1\}$. Also, $h \models (\phi, \alpha)$. By construction of h , $h^v = I$, so $\alpha \in I$. Moreover, for every $\beta \in I$, we know that $h_\beta^* = k_\beta^*$. Therefore $k_\alpha^* \equiv h_\alpha^*$. By Point 1 of Lemma 2, $h \models (\phi, \alpha)$ implies $h_\alpha^* \models \phi$. Then, $k_\alpha^* \models \phi$. As $k_\alpha^* \equiv t_\alpha^*$, we have that $t_\alpha^* \models \phi$. Again by Point 1 (of Lemma 2), $t_\alpha^* \models \phi$ iff $t \models (\phi, \alpha)$. Since $\alpha \geq \gamma$, $t \models (\phi, \gamma)$ by Point 2. The other direction can be proved similarly. \square

Theorem 8. Suppose \mathfrak{F} is an FO learning framework that is learnable with only membership queries. For all $p \in \mathbb{N}^+$, $\mathfrak{F}_\pi^p = (\mathcal{E}_\pi, \mathcal{L}_\pi^p)$ is learnable with only membership queries.

Proof. Let A be a learner for \mathfrak{F} and let $t \in \mathcal{L}_\pi^p$ be the target. For each $\alpha \in (0, 1]_p$, we run an instance of A , denoted A_α . Whenever A_α calls $\text{MQ}_{\mathfrak{F},t_\alpha^*}$ with ϕ as input, we call $\text{MQ}_{\mathfrak{F}_\pi, t}$ with (ϕ, α) as input. By Point 1 of Lemma 2, $\text{MQ}_{\mathfrak{F}, t_\alpha^*}(\phi) = \text{MQ}_{\mathfrak{F}_\pi, t}(\phi, \alpha)$. Since A is a learner for \mathfrak{F} , every A_α eventually

halts and outputs a hypothesis k_α^* such that $k_\alpha^* \equiv t_\alpha^*$. Since $t \in \mathcal{L}_\pi^p$, $t^v \subseteq (0, 1]_p$. By Lemma 7, $t \equiv \{(\phi, \alpha) \mid \phi \in k_\alpha^*, \alpha \in (0, 1]_p\}$. Thus, we can transfer learnability of \mathfrak{F} (with only membership queries) to \mathfrak{F}_π^p . \square

If, e.g., $\text{MQ}_{\mathfrak{F}_\pi, t}((\phi, 0.01)) = \text{'yes'}$, $\text{MQ}_{\mathfrak{F}_\pi, t}((\phi, 0.02)) = \text{'no'}$, and the precision of the target is 2, then $\text{val}(\phi, t) = 0.01$. So, knowing the precision is important for learning with membership queries only. If equivalence queries are allowed then a learner can build a hypothesis equivalent to the target *without knowing the precision* in advance by simply enumerating all possible hypothesis and asking them to the oracle, one by one (Theorem 9).

Theorem 9. *The possibilistic extension \mathfrak{F}_π of an FO learning framework \mathfrak{F} is learnable with only equivalence queries.*

If both membership and equivalence query oracles are available, learnability is guaranteed by the previous theorem.

Corollary 1. *Let \mathfrak{F} be an FO learning framework. \mathfrak{F} is learnable iff \mathfrak{F}_π is learnable.*

4 Polynomial Time Reduction

We now investigate whether results showing that an FO learning framework is in PTIMEL can be transferred to their possibilistic extensions and vice-versa. Theorem 10 shows the transferability of PTIMEL membership from the possibilistic extension \mathfrak{F}_π of an FO learning framework \mathfrak{F} to \mathfrak{F} .

Theorem 10. *Let \mathfrak{F} be an FO learning framework. If \mathfrak{F}_π is in PTIMEL then \mathfrak{F} is in PTIMEL.*

Proof. In our proof, we use the following claim.

Claim 10.1. *Let k be an FO KB and let t be the possibilistic KB $\{(\phi, 1) \mid \phi \in k\}$. For all (ϕ, α) , $k \models \phi$ iff $t \models (\phi, \alpha)$.*

Proof. If $t \models (\phi, \alpha)$, since $t^* \models t_\alpha^*$ and $k = t^*$, $k \models \phi$. If $k \models \phi$, by construction $t_1^* \models \phi$. By Point 1 of Lemma 2, $t_1^* \models \phi$ iff $t \models (\phi, 1)$, so, for all $\alpha \in (0, 1]$, $t \models (\phi, \alpha)$. \square

Let $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ and let $k \in \mathcal{L}$ be the target. Since \mathfrak{F}_π is in PTIMEL, there is a learner A_π for \mathfrak{F}_π . We start the execution of A_π that attempts to learn a hypothesis h equivalent to $t = \{(\phi, 1) \mid \phi \in k\}$. By Claim 10.1, for all $\alpha \in (0, 1]$, $\text{MQ}_{\mathfrak{F}_\pi, t}((\phi, \alpha)) = \text{MQ}_{\mathfrak{F}, k}(\phi)$. Also, we can simulate a call to $\text{EQ}_{\mathfrak{F}_\pi, t}$ with h as input by calling $\text{EQ}_{\mathfrak{F}, k}$ with h^* as input. By Claim 10.1, for all $\alpha \in (0, 1]$, $k \models \phi$ iff $t \models (\phi, \alpha)$, in particular, for $\alpha = 1$. By Remark 1, we can assume that all counterexamples returned by $\text{EQ}_{\mathfrak{F}, k}$ are positive. Whenever we receive a (positive) counterexample ϕ , we return $(\phi, 1)$ to A_π . Eventually, A_π will output a hypothesis $h \equiv t$ in polynomial time w.r.t. $|t|$ and the largest counterexample received so far. Clearly, h^* is as required. \square

By Theorem 11, the converse of Theorem 10 does not hold.

Theorem 11. *There exists an FO learning framework \mathfrak{F} such that \mathfrak{F} is in PTIMEL but $\mathfrak{F}_\pi = (\mathcal{E}_\pi, \mathcal{L}_\pi)$ is not in PTIMEL.*

Proof. Let $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ be an FO learning framework that is *not* in PTIMEL. Such \mathfrak{F} exists, one can consider, for instance, the \mathcal{EL} learning framework [Konev *et al.*, 2018, Theorem 68]². We use \mathfrak{F} to define the learning framework $\mathfrak{F}^\perp = (\mathcal{E}, \mathcal{L}^\perp)$ where $\mathcal{L}^\perp = \{h \cup \{\phi, \neg\phi\} \mid h \in \mathcal{L}\}$ for a fixed but arbitrary non-trivial FO formula ϕ . Even though \mathfrak{F} is not learnable in polynomial time, \mathfrak{F}^\perp is. The learner can learn any $l \in \mathcal{L}^\perp$ by returning the hypothesis $\{\perp\}$ (in constant time). Assume that $\mathfrak{F}_\pi^\perp = (\mathcal{E}_\pi, \mathcal{L}_\pi^\perp)$ is in PTIMEL. This means that for every target $l \in \mathcal{L}_\pi^\perp$ we can learn in polynomial time a hypothesis h such that $h \equiv l$. By construction, for every $t \in \mathcal{L}$ there is $l \in \mathcal{L}_\pi^\perp$ such that $t \equiv l_{\text{inc}(t)}^*$. By learning h such that $h \equiv l$ we have also learned a hypothesis h such that $h_{\text{inc}(h)}^* \equiv t$. By Theorem 10, $\mathfrak{F} \in \text{PTIMEL}$, which contradicts our assumption that this is not the case. Therefore we have found an FO learning framework \mathfrak{F}^\perp that is in PTIMEL but its possibilistic extension \mathfrak{F}_π^\perp is not in PTIMEL. \square

The FO learning framework \mathfrak{F}^\perp in the proof of Theorem 11 is not safe (see definition in Subsection 2.2) because, for $l \not\subseteq \{\phi, \neg\phi\}$ we have $l \in \mathcal{L}^\perp$ with $(l \setminus \{\phi, \neg\phi\}) \notin \mathcal{L}^\perp$. Intuitively, non-safe learning frameworks allow cases in which the target is easy to learn if we aim at learning the *whole* target, not a *subset* of it. In the following, we focus on FO learning frameworks that are safe³. The first transferability result we present is for the case in which the learner has access to only membership queries. To show the reduction, we use a procedure called FindValuation_t . This procedure takes as input a precision p and a formula ϕ and returns the highest valuation β with precision p of a formula ϕ entailed by the target t (or zero if it is not entailed). That is, β is such that $\beta =_p \text{val}(\phi, t)$. For any $\gamma \in [0, 1]_p$ the procedure can check if $t \models (\phi, \gamma)$ by calling the oracle $\text{MQ}_{\mathfrak{F}_\pi, t}$ with (ϕ, γ) as input. To compute β such that $\beta =_p \text{val}(\phi, t)$, FindValuation_t performs a binary search on $[0, 1]_p$. Lemma 12 states the correctness and the complexity of FindValuation_t .

Lemma 12. *Let $\mathfrak{F}_\pi = (\mathcal{E}_\pi, \mathcal{L}_\pi)$ be a possibilistic learning framework and let $t \in \mathcal{L}_\pi$ be the target. FindValuation_t , with input a precision $p \in \mathbb{N}^+$ and $\phi \in \mathcal{E}_\pi$, runs in polynomial time in p and $|\phi|$ and outputs β such that $\beta =_p \text{val}(\phi, t)$.*

Sketch. By Point 2 of Lemma 2, FindValuation_t can determine β such that $\beta =_p \text{val}(\phi, t)$ by performing a binary search on the interval of numbers $[0, 1]_p$. So the number of iterations is bounded by $\log_2(10^p + 1)$, which is polynomial in p . Each iteration can be performed in polynomial time in $|\phi|$ and p . \square

By Thm. 13, for safe FO learning frameworks, polynomial time results with only membership queries can be transferred to their possibilistic extensions if the precision of the target is known (by Thm. 6, we cannot remove this assumption).

Theorem 13. *Let \mathfrak{F} be a safe FO learning framework. For all $p \in \mathbb{N}^+$, when only membership queries can be asked, \mathfrak{F} is in PTIMEL iff \mathfrak{F}_π^p is in PTIMEL.*

²Non-polynomial query learnability is proved in [Konev *et al.*, 2018, Theorem 68], which implies non-polynomial time learnability.

³All learning from entailment results we found in the literature could be formulated in terms of safe learning frameworks.

Proof. To show the transferability of PTIMEL membership from \mathfrak{F} to \mathfrak{F}_π , we use the following claim.

Claim 13.1. *Assume $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ is safe and in PTIMEL with only membership queries. For every $p \in \mathbb{N}^+$ and framework $\mathfrak{F}_\pi^p = (\mathcal{E}_\pi, \mathcal{L}_\pi^p)$ with $t \in \mathcal{L}_\pi^p$, given a valuation α with $\text{prec}(\alpha) = p$, one can learn k_α^* such that $k_\alpha^* \equiv t_\alpha^*$ in time polynomial w.r.t. $|t|$ with only membership queries.*

Proof. We start the execution of a polynomial time learner A for \mathfrak{F} . Whenever A calls $\text{MQ}_{\mathfrak{F}, t_\alpha^*}$ with ϕ as input, we call $\text{MQ}_{\mathfrak{F}_\pi, t}$ with $(\phi, \alpha + 10^{-p})$ as input and we return the same answer to A . By Point 1 of Lemma 2, $\text{MQ}_{\mathfrak{F}, t_\alpha^*}(\phi) = \text{MQ}_{\mathfrak{F}_\pi, t}(\phi, \alpha + 10^{-p})$. Since \mathfrak{F} is safe, A will build a hypothesis k_α^* such that $k_\alpha^* \equiv t_\alpha^*$ in polynomial time w.r.t. $|t|$. \square

We set $\gamma := 0$ and $S := \emptyset$. By Claim 13.1 we can find in polynomial time w.r.t. $|t|$ a hypothesis k_γ^* such that $k_\gamma^* \equiv t_\gamma^*$. For every $\phi \in k_\gamma^*$, we run FindValuation_t with $p = \text{prec}(t)$ and ϕ as input to find $\text{val}(\phi, t)$. In this way, by Point 3 of Lemma 2 and Lemma 12, we identify in polynomial time w.r.t. $|t|$ some $\beta \in t^v \cup \{1\}$ such that $k_\gamma^* \equiv t_\beta^*$. We set $k_\beta^* := k_\gamma^*$ and add k_β^* to S . Then, we update γ to the value β and apply Claim 13.1 again. For every $\phi \in k_\gamma^*$, we run FindValuation_t again with $p = \text{prec}(t)$ and ϕ as input to find $\text{val}(\phi, t)$. We repeat this process until we find $k_\gamma^* \equiv \emptyset$ or $\gamma + 10^{-p} > 1$. Each time we run FindValuation_t , we identify a higher valuation in t^v . Therefore, this happens at most $|t^v|$ times. For all $\alpha \in t^v$, there is $k_\alpha^* \in S$ that satisfies $k_\alpha^* \equiv t_\alpha^*$, therefore, by Lemma 7,

$$h = \bigcup_{k_\alpha^* \in S} \{(\phi, \alpha) \mid \phi \in k_\alpha^*\}$$

is such that $h \equiv t$.

We now show the transferability of PTIMEL membership from \mathfrak{F}_π to \mathfrak{F} . Let $k \in \mathcal{L}$ be the target. We start the execution of a learner A_π for \mathfrak{F}_π that attempts to learn a hypothesis equivalent to $t = \{(\phi, 1) \mid \phi \in k\}$. By Claim 10.1 of Theorem 10, we can simulate a call to $\text{MQ}_{\mathfrak{F}_\pi, t}$ with input $(\phi, 1)$ by calling $\text{MQ}_{\mathfrak{F}, k}$ with ϕ as input and returning the same answer to A_π . A_π terminates in polynomial time w.r.t. $|t|$ with a hypothesis h such that $h \equiv t$. As $h^* \equiv t^* = k$, h^* is as required. \square

When we want to transfer learnability results from \mathfrak{F} to \mathfrak{F}_π it is important to learn one h_α such that $h_\alpha \equiv t_\alpha$ for each $\alpha \in t^v$, where t is the target (Example 14).

Example 14. *Let $t = \{(p \rightarrow q_1, 0.3), (p \rightarrow q_2, 0.7)\}$. We can use the polynomial time algorithm for propositional Horn [Frazier and Pitt, 1993] to learn a hypothesis $k^* = \{p \rightarrow (q_1 \wedge q_2)\} \equiv t^*$. However, if $h = \{(\phi, \text{val}(\phi, t)) \mid \phi \in k^*\}$ then $h = \{(p \rightarrow (q_1 \wedge q_2), 0.3)\} \not\equiv t$.*

A learner that has access to both membership and equivalence query oracle has a way of finding the precision of the target when it is unknown. With membership queries, we can use FindValuation_t to find the valuation of formulas up to a given precision. By Lemma 15, we can obtain useful information about the precision of the target with the counterexamples obtained after an equivalence query.

Lemma 15. *Assume $\mathfrak{F}_\pi = (\mathcal{E}_\pi, \mathcal{L}_\pi)$ is the possibilistic extension of a safe FO learning framework and $t \in \mathcal{L}_\pi$ is the target. Given $p \in \mathbb{N}^+$, one can determine that $p < \text{prec}(t)$ or compute $h \in \mathcal{L}_\pi$ such that $h \equiv t$, in polynomial time w.r.t. $|t|$, p , and the largest counterexample seen so far.*

Proof. In our proof, we use the following claims.

Claim 15.1. *Given $h \in \mathcal{L}_\pi$ such that $t \models h$, one can construct in polynomial time in $|h|$ some $h' \in \mathcal{L}_\pi$ such that $t \models h' \models h$ and, for all $(\phi, \alpha) \in h'$, $t \models (\phi, \alpha)$ and $\alpha =_{\text{prec}(h')} \text{val}(\phi, t)$.*

Proof. Let h' be the set of all (ϕ, β) such that $(\phi, \alpha) \in h$ and FindValuation_t returns β with ϕ and $\text{prec}(h)$ as input. As $t \models h$, by construction of h' , $t \models h' \models h$. By Lemma 12, h' can be constructed in polynomial time in $|h|$ and is as required. \square

Claim 15.2. *Let $h \in \mathcal{L}_\pi$ be such that, for all $(\phi, \alpha) \in h$, $t \models (\phi, \alpha)$ and $\alpha =_{\text{prec}(h)} \text{val}(\phi, t)$. If $\text{EQ}_{\mathfrak{F}_\pi, t}$ with input h returns (ϕ, α) then either we know that $\text{prec}(t) > \text{prec}(h)$ or $h_\beta^* \not\equiv \phi$ where $\beta =_{\text{prec}(h)} \text{val}(\phi, t)$.*

Proof. By Point 1 of Lemma 2, $h_\beta^* \models \phi$ iff $h \models (\phi, \beta)$. If $h \models (\phi, \beta)$ or $\beta = 0$ (note: β can be 0 because, e.g., $0.01 =_1 0$), then $\text{prec}(\text{val}(\phi, t)) > \text{prec}(h)$. By Point 3 of Lemma 2, $\text{val}(\phi, t) \in t^v \cup \{1\}$, so $\text{prec}(t) > \text{prec}(h)$. \square

By Remark 1, we can assume at all times in this proof that any hypothesis constructed is entailed by the target (possibilistic or not). Moreover, by Claim 15.1, we can assume that, for any target and hypothesis $t, h \in \mathcal{L}_\pi$, we have that, for all $(\phi, \alpha) \in h$, $t \models (\phi, \alpha)$ and $\alpha =_{\text{prec}(h)} \text{val}(\phi, t)$. So we can assume at all times in our proof that the hypothesis h we construct (Equation 1) satisfies the conditions of Claim 15.2.

Let A be a polynomial time learner⁴ for \mathfrak{F} . As in the proof of Theorem 13, we run multiple instances of A . We denote by R the set of instances of A . Each instance in R is denoted A_β and attempts to learn a hypothesis equivalent to t_β^* , where β is a valuation. We sometimes write A_β^n to indicate that the instance A_β has asked n equivalence queries so far. We denote by $k^{\beta, n}$ the hypothesis given as input by A_β^n when it asks its n -th equivalence query. For $n = 0$, we assume that $k^{\beta, 0} = \emptyset$.

Initially, $R := \{A_{10^{-p}}^0\}$. Whenever $A_\beta \in R$ asks a membership query with input $\phi \in \mathcal{E}$, by Point 1 of Lemma 2, we can simulate $\text{MQ}_{\mathfrak{F}, t_\beta^*}$ by calling $\text{MQ}_{\mathfrak{F}_\pi, t}$ with (ϕ, β) as input and returning the same answer to A_β . Let h_0 be $\{(\phi_\top, \alpha)\}$ where ϕ_\top is a tautology and α is a valuation with $\text{prec}(\alpha) = p$. Whenever $A_\beta^n \in R$ asks its n -th equivalence query, we leave A_β^n waiting in the query state (see description of a learning system in Subsection 2.2). When all $A_\alpha^m \in R$ are waiting in the query state, we create

$$h := \bigcup_{A_\alpha^m \in R} \{(\phi, \alpha) \mid \phi \in k^{\alpha, m}\} \cup h_0 \quad (1)$$

⁴Assume w.l.o.g. that A always eventually asks an equivalence query until it finds an equivalent hypothesis (but may execute other steps and ask membership queries between each equivalence query).

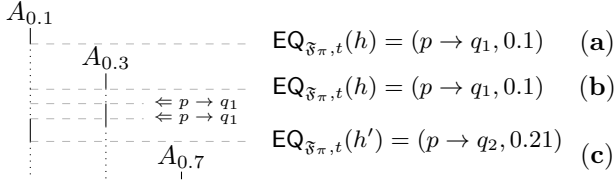


Figure 1: Multiple instances of algorithm A in Example 16. Time flows top-down. A dotted line means that the learner is waiting in query state, a continuous line means that the learner is running.

and call $\text{EQ}_{\mathfrak{F},t}$ with h as input (note: each instance $A_\alpha \in R$ may have asked a different number of equivalence queries when A_β^n asks its n -th equivalence query). If the answer is ‘yes’, we have computed h such that $h \equiv t$ and we are done. Upon receiving a (positive) counterexample (ϕ, γ) , we run FindValuation_t with ϕ and $\text{prec}(h)$ as input and compute a valuation β such that $\beta =_{\text{prec}(h)} \text{val}(\phi, t)$ (Lemma 12). If $A_\beta \notin R$, we start the execution of the instance A_β of algorithm A and add A_β to R . Otherwise, $A_\beta \in R$ and we check whether $k^{\beta,m} \models \phi$ (assume m is the number of equivalence queries posed so far by A_β). If $k^{\beta,m} \models \phi$ then, by Claim 15.2, we know that $\text{prec}(h) < \text{prec}(t)$ then we are done. If $k^{\beta,m} \not\models \phi$ then ϕ is a (positive) counterexample for $k^{\beta,m}$ and t_β^* . We return ϕ to every $A_\alpha^n \in R$ such that $\alpha \leq \beta$ and $k^{\alpha,m} \not\models \phi$ and these instances resume their executions. Observe that, since $h_0 \subseteq h$, by the construction of h , at all times $\text{prec}(h) = p$.

We now argue that this procedure terminates in polynomial time w.r.t. $|t|$, p , and the largest counterexample seen so far. Since there is only one instance A_β in R for each valuation β such that $\beta =_p \text{val}(\phi, t)$, by Point 3 of Lemma 2, we have that at all times $|R|$ is linear in $|t^v|$, which is bounded by $|t|$. By Lemma 12, whenever we run FindValuation_t to compute a valuation with ϕ and p as input, only polynomially many steps in $|\phi|$ and p are needed. Since \mathfrak{F} is safe and A is a polynomial time learner for \mathfrak{F} either we can determine that $p < \text{prec}(t)$ or each $A_\beta \in R$ terminates, in polynomial time in the size of t_β^* and the largest counterexample seen so far, and outputs $k^{\beta,n} = h_\beta^*$ such that $h_\beta^* \equiv t_\beta^*$. In this case, by Lemma 7, $h \equiv t$ and the process terminates. \square

The constructive proof of Lemma 15 delineates the steps made in Example 16 where the precision of the target is 1.

Example 16. Let $\mathfrak{F} = (\mathcal{E}, \mathcal{L})$ be the safe learning framework where \mathcal{L} is the set of all propositional Horn KBs and \mathcal{E} is the set of all (propositional) Horn clauses. Let $t \in \mathcal{L}_\pi$ and A be, respectively, the target and the learner of Example 14. Following our argument in Lemma 15, we start an instance $A_{0.1}$ of A . When $A_{0.1}$ is waiting in the query state, we build $h = \{(\phi_\top, 0.1)\}$ (Equation 1) and call $\text{EQ}_{\mathfrak{F},t}$ with h as input (Point (a) in Figure 1). Assume we receive the positive counterexample $(p \rightarrow q_1, 0.1)$. We run FindValuation_t with 1 and $p \rightarrow q_1$ as input, which computes $\text{val}(p \rightarrow q_1, t) = 0.3$. Since $A_{0.3} \notin R$, we start $A_{0.3}$. When all learners are waiting in the query state, we call again $\text{EQ}_{\mathfrak{F},t}$ with h as input (Point (b) in Figure 1). At this point, $R = \{A_{0.1}, A_{0.3}\}$.

Assume we receive $(p \rightarrow q_1, 0.1)$ again. We have that $\text{val}(p \rightarrow q_1, t) = 0.3$ and $A_{0.3} \in R$. Since $k^{0.3,1} \not\models p \rightarrow q_1$

and $k^{0.1,1} \not\models p \rightarrow q_1$, we return $p \rightarrow q_1$ to both $A_{0.1}$ and $A_{0.3}$ and they resume their executions. All learners will eventually be waiting in query state. When this happens we call $\text{EQ}_{\mathfrak{F},t}$ with $h' = \{(\phi_\top, 0.1), (p \rightarrow q_1, 0.1), (p \rightarrow q_1, 0.3)\}$ as input.

Assume the response is $(p \rightarrow q_2, 0.21)$. We run FindValuation_t with 1 and $p \rightarrow q_2$ as input, which returns $\text{val}(p \rightarrow q_2, t) = 0.7$. As before, we start $A_{0.7}$ (Point (c) in Figure 1) and add it to R . When all learners are waiting again we call $\text{EQ}_{\mathfrak{F},t}$ with h' as input. Assume we receive $(p \rightarrow q_2, 0.1)$. We then send $p \rightarrow q_2$ to every learner in R . Next time we call $\text{EQ}_{\mathfrak{F},t}$, with $h' \cup \{(p \rightarrow q_2, 0.7), (p \rightarrow q_2, 0.3), (p \rightarrow q_2, 0.1)\}$ as input. The answer is ‘yes’ and we are done. \triangleleft

A direct consequence of Lemma 15 is Theorem 17.

Theorem 17. *For every safe FO learning frameworks \mathfrak{F} we have, \mathfrak{F} is in PTIMEL iff \mathfrak{F}_π is in PTIMEL.*

Proof. One direction holds by Theorem 10. We prove the other direction. Let \mathfrak{F} be a safe FO learning framework in PTIMEL and let $\mathfrak{F}_\pi = (\mathcal{E}_\pi, \mathcal{L}_\pi)$ be its possibilistic extension. Consider a learner that initially estimates precision p of the target $t \in \mathcal{L}_\pi$ to be 1. Using Lemma 15, we can assume that this learner can either determine that $p < \text{prec}(t)$ or find a hypothesis h such that $h \equiv t$, in time polynomial with respect to $|t|$, p and the largest counterexample seen so far. In the former case, this learner sets the estimated precision p of the target to $p + 1$. This happens at most $\text{prec}(t)$ times, which is bounded by $|t|$. As a consequence, \mathfrak{F}_π is in PTIMEL. \square

We end this section recalling that our results can be transferred to the PAC model [Valiant, 1984] extended with membership queries (Theorem 18).

Theorem 18 ([Angluin, 1988; Mohri *et al.*, 2012]). *Let PTIMEPL be the class of all learning frameworks that are PAC learnable with membership queries in polynomial time. Then, PTIMEL \subseteq PTIMEPL.*

By Theorems 17 and 18, the following holds.

Corollary 2. *For all safe FO learning frameworks \mathfrak{F} , if $\mathfrak{F} \in \text{PTIMEL}$ then $\mathfrak{F}_\pi \in \text{PTIMEPL}$.*

5 Conclusion

Uncertainty is widespread in learning processes. Among different uncertainty formalisms, possibilistic logic stands out because of its ability to express preferences among worlds and model ignorance. We presented the first study on the exact (polynomial) learnability of possibilistic theories. It follows from our results that various algorithms designed for exact learning fragments of first-order logic can be adapted to learn their possibilistic extensions. We leave open the problem of polynomial time transferability with only equivalence queries.

Acknowledgements

We are supported by the University of Bergen. We would like to thank Andrea Mazzullo for joining initial discussions.

References

- [Agarwal and Nayal, 2015] Parul Agarwal and Dr. H. S. Nayal. Possibility theory versus probability theory in fuzzy measure theory. *International Journal of Engineering Research and Applications*, 5(5):37–43, 2015.
- [Angluin *et al.*, 1992] Dana Angluin, Michael Frazier, and Leonard Pitt. Learning conjunctions of horn clauses. *Machine Learning*, 9:147–164, 1992.
- [Angluin, 1988] Dana Angluin. Queries and concept learning. *Machine Learning*, 2(4):319–342, 1988.
- [Arimura, 1997] Hiroki Arimura. Learning acyclic first-order Horn sentences from entailment. In *International Workshop on Algorithmic Learning Theory*, pages 432–445, 1997.
- [Benferhat *et al.*, 1992] Salem Benferhat, Didier Dubois, and Henri Prade. Representing default rules in possibilistic logic. In *KR*, page 673–684. Morgan Kaufmann Publishers Inc., 1992.
- [Didier Dubois and Prade, 1994] Jérôme Lang Didier Dubois and Henri Prade. *Possibilistic Logic*, page 439–513. Oxford University Press, Inc., USA, 1994.
- [Djouadi *et al.*, 2010] Yassine Djouadi, Didier Dubois, and Henri Prade. Possibility theory and formal concept analysis: Context decomposition and uncertainty handling. pages 260–269, 06 2010.
- [Dubois and Prade, 1993] Didier Dubois and Henri Prade. Fuzzy sets and probability: misunderstandings, bridges and gaps. In *IEEE International Conference on Fuzzy Systems*, volume 2, pages 1059–1068, 1993.
- [Dubois and Prade, 2001] Didier Dubois and Henri Prade. Possibility theory, probability theory and multiple-valued logics: A clarification. *Ann. Math. Artif. Intell.*, 32(1-4):35–66, 2001.
- [Dubois and Prade, 2015] Didier Dubois and Henri Prade. Possibility theory and its applications: Where do we stand? In *Handbook of Computational Intelligence*, Springer Handbooks, pages 31–60. Springer, 2015.
- [Frazier and Pitt, 1993] Michael Frazier and Leonard Pitt. Learning from entailment: An application to propositional Horn sentences. In *ICML*, pages 120–127, 1993.
- [Hermo and Ozaki, 2020] Montserrat Hermo and Ana Ozaki. Exact learning: On the boundary between horn and CNF. *TOCT*, 12(1):4:1–4:25, 2020.
- [Konev *et al.*, 2018] Boris Konev, Carsten Lutz, Ana Ozaki, and Frank Wolter. Exact Learning of Lightweight Description Logic Ontologies. *Journal of Machine Learning Research*, 18(201):1–63, 2018.
- [Kuzelka *et al.*, 2015] Ondrej Kuzelka, Jesse Davis, and Steven Schockaert. Encoding markov logic networks in possibilistic logic. In *UAI*, pages 454–463. AUAI Press, 2015.
- [Kuzelka *et al.*, 2016] Ondrej Kuzelka, Jesse Davis, and Steven Schockaert. Learning possibilistic logic theories from default rules. In *IJCAI*, pages 1167–1173, 2016.
- [Kuzelka *et al.*, 2017] Ondrej Kuzelka, Jesse Davis, and Steven Schockaert. Induction of interpretable possibilistic logic theories from relational data. In *IJCAI*, pages 1153–1159, 2017.
- [Lang, 2000] Jérôme Lang. *Possibilistic Logic: Complexity and Algorithms*, pages 179–220. Springer Netherlands, 2000.
- [Mohri *et al.*, 2012] Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar. *Foundations of Machine Learning*. Adaptive computation and machine learning. MIT Press, 2012.
- [Parsons and Hunter, 1998] Simon Parsons and Anthony Hunter. *A Review of Uncertainty Handling Formalisms*, pages 8–37. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998.
- [Prade and Serrurier, 2008] Henri Prade and Mathieu Serrurier. Bipolar version space learning. *International Journal of Intelligent Systems*, 23(10):1135–1152, 2008.
- [Reddy and Tadepalli, 1998] Chandra Reddy and Prasad Tadepalli. Learning first-order acyclic Horn programs from entailment. *ILP*, pages 23–37, 1998.
- [Serrurier and Prade, 2007] Mathieu Serrurier and Henri Prade. Introducing possibilistic logic in ILP for dealing with exceptions. *Artif. Intell.*, 171(16-17):939–950, 2007.
- [Valiant, 1984] Leslie Gabriel Valiant. A theory of the learnable. *Commun. ACM*, 27(11):1134–1142, 1984.
- [Watanabe, 1990] Osamu Watanabe. A formal study of learning via queries. In Michael S. Paterson, editor, *Automata, Languages and Programming*, pages 139–152, Berlin, Heidelberg, 1990. Springer Berlin Heidelberg.