

1 **Reviewer 1** Sorry for the confusion. The argument is roughly that because Theorem 4 establishes that  $\text{GND}(\alpha, 1)$  is  
2 satisfiable if and only if  $\alpha$  is in the countably infinite domain, and Lemma 5 establishes that the introduction of extra  
3 names in  $\text{GND}(\alpha, z)$  preserves satisfiability, we obtain satisfiability under the larger, common subset of names used in  
4  $\text{GND}^-$  as well. Please see the Appendix of Ref. 1 for a formal proof.

5 For your second question, note that for our purposes, it is enough to only consider masking functions that mask all but a  
6 finite subset of the domain, thus producing finite-size partial models. Thus, we can take both our masking function and  
7 the “application function”  $\text{app}(M, \Theta)$  to have a countable range, not something with continuum (or larger) cardinality.  
8 In particular, then, we get measurability: these discrete output sets can be defined by a countable union over the resulting  
9 finite partial examples. The preimage of a single partial model in turn will be a measurable set for  $M$  and  $\Theta$ , given that  
10  $M$  is a measurable function: by definition, the preimage for a measurable  $M$  of that partial model is a measurable set.

11 **Reviewer 2** The representation language  $\text{proper}^+$  that we use is emerging as a popular representation language.  
12 FOL with universal quantifiers is widely used to express inductive properties in mathematics but also to represent  
13 social networks and graphs. For computability results, usually the finite domain assumption is made, but interestingly,  
14  $\text{proper}^+$  seems to allow us to go beyond the closed-world assumption. (And unlike description logics, arity restrictions  
15 are also not needed.) We note that beyond the fact that  $\text{proper}^+$  extends incomplete databases (L104–107), for example  
16 (Liu and Lakemeyer, 2009) show how to represent a certain family of “local” action models for planning within the  
17 fragment of  $\text{proper}^+$  for which Theorems 14/16 give polynomial-time reasoning. There is also a variant for epistemic  
18 planning (Muise et al. 2015) where one reasons about the mental states of other agents, and we expect analogous  
19 extensions of our work may contribute to that direction too. In particular, our approach applies to infinite domains, or  
20 even simply large domains without resorting to directly considering all groundings of the atomic formulas, in contrast  
21 to Juba’s work. Note that even in moderate size, finite domains, the number of groundings grows exponentially with the  
22 arity of the formulas under consideration, and thus quickly grows infeasible to represent as a propositional formula  
23 (which is required for Juba’s approach).

24 Implicit learning works because the partial models themselves compactly encode all of the rules that could be learned  
25 from those models. So instead of trying to learn a large set of rules from the models and hoping that these rules will  
26 permit us to derive the desired conclusions, we use the models directly to answer queries.

27 **Reviewer 3** Sorry for the terse exposition. The *universal closure* is the result of placing a universal quantifier on each  
28 free variable appearing in the formula.  $\supset$  (L81) denotes implication. Maximality (L107) refers to the database case, in  
29 which every true literal is included; by contrast we have a set of clauses that may not specify all of the true literals.  
30  $e\theta$  (L111) indeed refers to applying the substitution  $\theta$  to  $e$ .  $z$  (L116) is the rank, which yes is an integer. We use  $\mathbb{N}$   
31 as the set of names for convenience, but it is not important here that  $z$  could be interpreted as a name. In Proposition  
32 8 (L175) we simply take a union bound over the error events which have respective probabilities  $\epsilon_i$ . Note that  $1 - \epsilon'$   
33 validity only requires that the total probability of the error events is *at most*  $\epsilon'$ . The union bound applies to any set of  
34 events and in particular does not require independence. The actual guarantees of the informal discussion on L220 are  
35 formalized in Theorem 13. There is not a requirement on a distribution of queries. Rather, what we promise is that I:  
36 we will not (significantly) overestimate the validity of a query and II: we guarantee that our estimate of the validity is  
37 (approximately) at least the probability that some suitable implicit KB  $\mathcal{I}$  is witnessed.

38 Learning from entailment is pretty different from what we seek here: it asks us to produce a set of formulas  $H$  that  
39 replicates a desired set of entailment judgments, e.g., that  $\phi_1$  is entailed but  $\phi_2$  is not, etc. Our task formulation is  
40 much closer to learning from interpretations in ILP, where our partial models are partial interpretations. In that task,  
41 one is given a set of background knowledge formulas  $B$  and a set of models  $x_1, x_2, \dots$  and seeks an additional set  
42 of background knowledge  $H$  such that  $B \wedge H$  is consistent with the given models. One could subsequently answer  
43 entailment queries against  $B \wedge H$ . The difference is first that ILP only seeks an  $H$  that is consistent with the examples,  
44 and does not seek to analyze the degree to which the resulting formulas capture an unknown, ground-truth process that  
45 produced the example models  $x_1, x_2, \dots$ . In particular, there is no sense in which the resulting judgments  $B \wedge H \models \alpha$   
46 are “correct” or “incorrect” in ILP, unless we use a “closed-world” assumption or something similar, that leads the set  
47 of models to *define* a single “correct”  $H$ . Even so, in practice, ILP often requires significant restrictions on the set of  
48 clauses permitted in  $H$  to ensure that there is a finite Herbrand base of atoms to search through. We will include a  
49 discussion on this in the paper.

50 The time complexity bound is good in the sense that for fixed approximation and confidence parameters  $\gamma$  and  $\delta$ , the  
51 time complexity of querying the implicit KB is equivalent to a constant number of queries for an explicit KB (cf.  
52 Theorem 14). So it remains tractable.

53 **References** (a) Y. Liu and G. Lakemeyer. On first-order definability and computability of progression for local-effect actions and  
54 beyond. In Proc. IJCAI, pages 860–866, 2009. (b) C. J. Muise, V. Belle, P. Felli, S. A. McIlraith, T. Miller, A. R. Pearce, and L.  
55 Sonenberg. Planning over multi-agent epistemic states: A classical planning approach. In AAAI, 2015.