Pair-Valued Deontic Semantics and Rule-Regularised Learning from Conflicting Norms

Abstract

We develop a finite-domain, pair-valued deontic semantics for conflicting normative information. Formulas have separate support and opposition degrees at each world of a frame with graded accessibility; first-order quantifiers use finite-domain min/max aggregation and propositional formulas use strong Łukasiewicz connectives. We prove classical Kripke agreement on Boolean valuations and crisp frames, non-explosion, modal monotonicity, constant-domain quantifier commutation, sensitivity and evaluation bounds, and a crisp-frame containment result: positive inconsistency in a formula requires an inconsistent atom at a reachable world. A synthetic learning study tests only the single-ideal-world reduction, comparing data-only training, rule constraints with a consistency regulariser, and rule constraints without that regulariser. Rule constraints improve accuracy on a derived norm relative to data-only training. Without the consistency regulariser, source-norm accuracy and conflict detection improve when conflicts are frequent, but derived-norm accuracy remains lower than with the regulariser. The study does not test learned accessibility, multi-world modal learning, general first-order inference or source-provenance explanations. Code and per-seed summary metrics accompany the paper.

Keywords: paraconsistent logic; deontic logic; neurosymbolic AI; first-order logic; normative conflicts; rule-regularised learning

1. Introduction

Normative systems are rarely consistent. A compliance officer reads a data-retention rule from one regulator that requires records to be kept, and a deletion rule from another that requires them to be erased; a robot receives an instruction from its owner and a constraint from its manufacturer that cannot both be met; a clinical protocol and a hospital policy disagree. Philosophers of law and ethics have long treated such normative conflicts as real, and not as mistakes to be explained away.[1] Standard deontic logic (SDL), however, cannot accommodate them without inconsistency. Its axiom D states that whatever is obligatory is permitted, so that O p and O¬p together are inconsistent, and in a classical setting an inconsistent set of premises entails every formula.[2] For a reasoner that must act on a

[1]Goble, L. (2009). Normative conflicts and the logic of 'ought'. *Noûs, 43*(3), 450–489. https://doi.org/10.1111/j.1468-0068.2009.00714.x discusses normative conflicts and develops deontic logics that admit them.
[2]von Wright, G. H. (1951). Deontic logic. *Mind, 60*(237), 1–15. SDL is the normal modal logic KD read deontically; its Kripke semantics (Kripke, S. A. (1963). Semantical analysis of modal logic I: Normal modal propositional calculi.

normative specification, that behaviour is useless: the conflict that most needs to be reported is exactly the one that destroys the ability to report anything.

The problem becomes sharper when normative knowledge is combined with learning. Neurosymbolic AI aims to integrate neural learning with symbolic knowledge so that systems are accurate, data-efficient and explainable.[3] Frameworks such as Logic Tensor Networks turn first-order formulas into differentiable constraints over real-valued truth degrees,[4] and probabilistic logic programming has been coupled with neural predicates.[5] If one scalar output is trained by cross-entropy on one assertion of p and one of ¬p, its optimum is one half; that particular representation cannot distinguish disagreement from an uninformative degree of one half. This does not rule out other scalar encodings or losses. What is needed is a logic whose values distinguish *no information*, *information for*, *information against* and *information both for and against*, whose modal operators express obligation and permission, and whose semantics remains differentiable.

Such values exist in the paraconsistent tradition. Belnap's four-valued logic assigns to a sentence any subset of {told true, told false},[6] modal logics with these truth values have been studied,[7] and graded paraconsistent modal logics with two real-valued coordinates for positive and negative information exist as well.[8] On the neural side, Modal Logic Neural Networks carry truth as an interval of bounds and include deontic readings and a deontic experiment (Section 2.1). This paper builds on that line and does not claim the propositional and modal clauses as new: Section 2.4 shows that, under the change of variables $L = t$, $U = 1 - f$, our clauses are the componentwise Łukasiewicz interval operators. What we add is (i) a normative reading of the pair that separates the pair of a formula $O\alpha$ from the status of α (Section 3.3), (ii) finite-domain first-order quantification and a set of stated and proved

Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 9(5–6), 67–96. https://doi.org/10.1002/malq.19630090502) interprets O as truth in all ideal, or deontically accessible, worlds.

[3]Garcez, A. d., & Lamb, L. C. (2023). Neurosymbolic AI: The 3rd wave. *Artificial Intelligence Review, 56*(11), 12387–12406. https://doi.org/10.1007/s10462-023-10448-w

[4]Badreddine, S., d'Avila Garcez, A., Serafini, L., & Spranger, M. (2022). Logic tensor networks. *Artificial Intelligence, 303*, Article 103649. https://doi.org/10.1016/j.artint.2021.103649

[5]Manhaeve, R., Dumančić, S., Kimmig, A., Demeester, T., & De Raedt, L. (2021). Neural probabilistic logic programming in DeepProbLog. *Artificial Intelligence, 298*, Article 103504. https://doi.org/10.1016/j.artint.2021.103504

[6]Belnap, N. D. (1977). A useful four-valued logic. In J. M. Dunn & G. Epstein (Eds.), *Modern uses of multiple-valued logic* (pp. 5–37). Springer Netherlands. https://doi.org/10.1007/978-94-010-1161-7_2 Dunn, J. M. (1976). Intuitive semantics for first-degree entailments and 'coupled trees'. *Philosophical Studies, 29*(3), 149–168.

[7]Odintsov, S. P., & Wansing, H. (2010). Modal logics with Belnapian truth values. *Journal of Applied Non-Classical Logics, 20*(3), 279–304. https://doi.org/10.3166/jancl.20.279-304

[8]Bílková, M., Frittella, S., & Kozhemiachenko, D. (2022). *Paraconsistent Gödel modal logic*. arXiv:2203.01237. https://arxiv.org/abs/2203.01237 The paper works with dual valuations into [0,1] on crisp modal frames and a Gödel-style algebra.

properties, and (iii) a controlled learning study of what rule constraints and a consistency penalty do when the training sources conflict.

Contributions. (i) We define PDL-$\mathbb{R}$, a pair-valued deontic semantics whose formulas take values in $[0,1]^2$ at each world of a frame with real-valued accessibility, with first-order quantification over a finite domain (Section 3). (ii) We state and prove ten results (Section 4): conservativity over classical Kripke evaluation on Boolean valuations, modal duality, modal monotonicity under truth-order entailment, non-explosion, failure of disjunctive syllogism for the material conditional, the Barcan formula and its converse, bounded sensitivity, an evaluation-cost bound, *containment* for crisp accessibility, and a proposition on scalar encodings. (iii) We describe how the semantics gives training signals (Section 5) and work through two examples (Section 6). (iv) We run a controlled study of a single-ideal-world reduction, in which a network learns norm statuses from two sources that overlap and conflict, and compare three loss configurations (Section 7). All numbers reported in this paper come from code supplied with the submission.

Scope. The main contribution is a formal semantics and its proved properties. The learning experiment is a controlled illustration of a single-ideal-world reduction, not an evaluation of the multi-world or quantified architecture. It uses synthetic normative sources with known ground truth and does not model a particular legal system or regulation. Sections 7.4 and 8 delimit what the results support.

2. Background and related work

2.1 Neurosymbolic learning with first-order and modal logic

Real Logic, the language of Logic Tensor Networks, grounds first-order symbols in real vectors and interprets connectives and quantifiers with fuzzy operators, so that the degree of satisfaction of a knowledge base is a differentiable function of the network parameters.[9] Logical Neural Networks give each formula a lower and an upper bound on its truth value and a weighted real-valued logic in which bounds are propagated in both directions.[10] Modal Logic Neural Networks extend this line to Kripke semantics: worlds, valuation and accessibility relation are all part of a single differentiable forward pass, and truth is carried as a bounded interval that lets the network represent inconsistent information; the current

[9]Badreddine et al. (2022), cited above. Serafini and Garcez's original formulation appeared as arXiv:1606.04422 (2016); we cite the journal version.
[10]Riegel, R., Gray, A., Luus, F., Khan, N., Makondo, N., Akhalwaya, I. Y., Qian, H., Fagin, R., Barahona, F., Sharma, U., Ikbal, S., Karanam, H., Neelam, S., Likhyani, A., & Srivastava, S. (2020). *Logical neural networks*. arXiv:2006.13155. https://arxiv.org/abs/2006.13155

version of that work also gives a deontic reading of the modalities and a deontic experiment.[11] Earlier work on connectionist modal logic showed how modal reasoning can be embedded in neural networks in the neural-symbolic tradition.[12] Temporal extensions of logic tensor networks to first-order linear temporal logic have also been proposed.[13] Of these, only the modal network work treats obligation.

2.2 Paraconsistent and many-valued logics

A logic is paraconsistent if its consequence relation does not trivialise in the presence of contradictions. Da Costa built formal systems with this property,[14] Priest proposed the Logic of Paradox, in which a sentence can be both true and false,[15] and Logics of Formal Inconsistency internalise a notion of consistency in the object language.[16] Belnap and Dunn's four values combine the two ideas: a value records whether there is evidence for a sentence and whether there is evidence against it, independently.[17] Our semantics is the unit-square analogue: the four corners are Belnap's values, and interior points express partial evidence. We use the Łukasiewicz t-norm and t-conorm, whose properties are treated in Hájek's monograph.[18]

2.3 Deontic logic, its conflicts and its paradoxes

SDL is sound and complete for serial Kripke frames. Two families of difficulty are well known. The first concerns conflicts: if an agent faces obligations O p and O¬p from independent sources, SDL derives O q for every q. The second concerns contrary-to-duty obligations. Chisholm's set of four sentences, an obligation to go, an obligation to tell if one goes, an obligation not to tell if one does not go, and the fact that one does not go, appears consistent and non-redundant yet cannot be expressed in SDL without collapse or

---

[11]Sulc, A., & Naddour, N. (2025). *Modal logic neural networks*. arXiv:2512.03491 (version 3 read for this revision; Sections 3.2 and 5.2 treat the deontic reading and a turbofan-engine safety experiment with deontic predicates). https://arxiv.org/abs/2512.03491 This is the work closest to ours; Section 2.4 compares the two. We do not compare against their implementation.

[12]Garcez, A. S. d'Avila, Lamb, L. C., & Gabbay, D. M. (2007). Connectionist modal logic: Representing modalities in neural networks. *Theoretical Computer Science, 371*(1–2), 34–53. Neural-symbolic learning is also treated in the monograph by Garcez, Broda, and Gabbay (2002), *Neural-symbolic learning systems: Foundations and applications*, Springer.

[13]Boscarato, L., Donadello, I., Artale, A., Montali, M., & Maggi, F. M. (2026). *First-order temporal logic tensor networks*. arXiv:2606.29972. https://arxiv.org/abs/2606.29972

[14]da Costa, N. C. A. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic, 15*(4), 497–510. https://doi.org/10.1305/ndjfl/1093891487

[15]Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic, 8*(1), 219–241.

[16]Carnielli, W., Coniglio, M. E., & Marcos, J. (2007). Logics of formal inconsistency. In D. M. Gabbay & F. Guenthner (Eds.), *Handbook of philosophical logic* (Vol. 14, pp. 1–93). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6324-4_1

[17]Belnap (1977) and Dunn (1976), cited above. Generalisations to arbitrary lattices of truth values, including for logic programming, are in Kifer, M., & Subrahmanian, V. S. (1992). Theory of generalized annotated logic programming and its applications. *The Journal of Logic Programming, 12*(4), 335–367.

[18]Hájek, P. (1998). *Metamathematics of fuzzy logic*. Kluwer Academic Publishers.

redundancy.[19] Goble's work develops logics of 'ought' that admit conflicts.[20] Our contribution to this line is not a new solution to the paradoxes but a semantics in which both problems can be posed inside a learning system, with values in the unit interval, and in which we can prove what is and is not contained.

2.4 Positioning and comparison

The closest neural work already handles inconsistency and gives deontic readings and an experiment; related logical work gives graded paraconsistent modalities without a learning study. Table 1 compares the representational choices of PDL-$\mathbb{R}$, scalar fuzzy logic-tensor approaches and interval-bound modal networks. Our mathematical novelty is not the componentwise propositional or modal operators, which are related by the coordinate transformation below, but the explicit deontic status distinction, finite-domain quantification, stated guarantees and controlled test of rule and consistency losses.

Dimension	**Scalar fuzzy (LTN-style)**	**Interval bounds (MLNN)**	**PDL-$\mathbb{R}$ (this paper)**
Truth values	one degree in [0,1]	interval [L,U] of bounds	pair (support t, opposition f) in $[0,1]^2$
Conflicting evidence	compromise degree (Proposition 10)	crossed bounds (L > U)	glut: t + f > 1, reported as B
Modalities	none in the base framework	Kripke modalities, with deontic reading	O and P over graded accessibility R
Accessibility	–	learnable	learnable in the architecture; fixed and crisp in the experiments
Quantifiers	aggregated over groundings	–	min/max over a finite constant domain
Evaluated here	–	not evaluated by us	synthetic study of a single-ideal-world reduction[21]

Table 1. Design dimensions. The entries for interval bounds are taken from the abstract and Sections 3.2 and 5.2 of the current version of the modal-network paper; we do not claim to have tested its implementation.

Relation to interval bounds. Under the change of variables L = t and U = 1 − f, the negation swap becomes the interval negation (L′, U′) = (1 − U, 1 − L); the strong conjunction becomes $T(L_1,L_2)$ on lower bounds and $T(U_1,U_2)$ on upper bounds, because $1 - \min(1, f_1+f_2) = \max(0, U_1+U_2-1)$; the strong disjunction becomes S on both; and the clauses for O act on both bounds as the Łukasiewicz implication R → ·, because $1 - \mathbf{T}(R, 1-U) = \mathbf{S}(1-R, U)$. A glut t

[19]Chisholm, R. M. (1963). Contrary-to-duty imperatives and deontic logic. *Analysis, 24*(2), 33–36.
[20]Goble (2009), cited above.
[21]The scope of the study is stated in Section 7.4: it evaluates a reduction of the framework and not the multi-world, quantified architecture.

+ f > 1 is exactly a crossed interval L > U. The propositional and modal clauses of PDL-$\mathbb{R}$ are therefore the componentwise interval operators, and we do not claim them as new. The two presentations differ in interpretation (evidence for and against, not bounds on one hidden truth value), in designated values, in the treatment of quantifiers, and in what is proved and tested.

3. The logic PDL-$\mathbb{R}$

3.1 Language

Let Π be a set of predicate symbols with arities and let Var be a set of variables. Terms are variables or constants from a finite domain $\mathbb{D}$. Formulas are given by

$$\varphi ::= \pi(t_1,\ldots,t_n) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \forall x\, \varphi \mid \exists x\, \varphi \mid O\,\varphi \mid P\,\varphi$$

with $\varphi \rightarrow \psi$ defined as $\neg\varphi \vee \psi$ and Forb φ (forbidden) defined as $O\neg\varphi$. Deontic status is given by the pattern of $O\varphi$ and $O\neg\varphi$. The connectives $\wedge$ and $\vee$ are the *strong* Łukasiewicz connectives: they are not idempotent away from the four corners (for p = (0.5, 0), the value of $p \wedge p$ is (0, 0)). No assumption of consistency or completeness is built into the syntax. The domain $\mathbb{D}$ is non-empty, and each constant denotes an element of $\mathbb{D}$.

3.2 Models and values

A *model* is M = (W, R, $\mathbb{D}$, V) where W is a finite non-empty set of worlds, R: W×W → [0,1] is a *deontic accessibility* function (R(w,v) is the degree to which v is an ideal alternative to w), $\mathbb{D}$ is the finite domain, and V assigns to every world w, predicate symbol π of arity n and tuple $d \in \mathbb{D}^n$ a pair $V(w,\pi,d) = (t,f) \in [0,1]^2$. We call t the *support* and f the *opposition*. The four corners are

$$T = (1,0)\ \text{true only};\quad F = (0,1)\ \text{false only};\quad B = (1,1)\ \text{both};\quad N = (0,0)\ \text{neither.}$$

We write $[\![\varphi]\!]_w = (t,f)$ for the value of φ at w under an assignment of variables that we leave implicit. Let $\mathbf{T}(a,b) = \max(0, a+b-1)$ and $\mathbf{S}(a,b) = \min(1, a+b)$ be the Łukasiewicz t-norm and t-conorm. The clauses are:

$$[\![\neg\varphi]\!]_w = (f_\varphi , t_\varphi)$$
$$[\![\varphi\wedge\psi]\!]_w = (\mathbf{T}(t_\varphi,t_\psi) , \mathbf{S}(f_\varphi,f_\psi))$$
$$[\![\varphi\vee\psi]\!]_w = (\mathbf{S}(t_\varphi,t_\psi) , \mathbf{T}(f_\varphi,f_\psi))$$
$$[\![\forall x\varphi]\!]_w = (\min_{d\in\mathbb{D}} t_\varphi[d] , \max_{d\in\mathbb{D}} f_\varphi[d])$$
$$[\![\exists x\varphi]\!]_w = (\max_{d\in\mathbb{D}} t_\varphi[d] , \min_{d\in\mathbb{D}} f_\varphi[d])$$
$$[\![O\varphi]\!]_w = (\min_{v\in W} \mathbf{S}(1-R(w,v), t_\varphi(v)) , \max_{v\in W} \mathbf{T}(R(w,v), f_\varphi(v)))$$
$$[\![P\varphi]\!]_w = [\![\neg O\neg\varphi]\!]_w = (\max_{v\in W} \mathbf{T}(R(w,v), t_\varphi(v)) , \min_{v\in W} \mathbf{S}(1-R(w,v), f_\varphi(v)))$$

The support of $O\varphi$ is the smallest degree to which every ideal alternative v supports φ, discounted by how ideal v is: $\mathbf{S}(1-R, t) = \min(1, 1-R+t)$ is the Łukasiewicz implication R →

t, so that an alternative with R = 0 imposes no requirement. The opposition of Oφ is the largest degree to which some ideal alternative opposes φ. When R takes only the values 0 and 1 these are the usual clauses, 'all successors' for support and 'some successor' for opposition.[22]

Designation and consequence. A value is *designated* if its support equals 1. A set of formulas Γ *entails* φ, written Γ ⊨ φ, if in every model, at every world and under every assignment at which all members of Γ are designated, φ is designated. We also use a stronger relation between formulas: φ ⩽ ψ ('φ entails ψ in the truth order') if in every model, at every world and under every assignment t_φ ≤ t_ψ and f_ψ ≤ f_φ. This is the semantic form of first-degree entailment.[23] Note that φ ⩽ ψ implies φ ⊨ ψ, but not conversely.

3.3 Deontic reading: the pair of Oα and the status of α

Two different pairs must be kept apart. The *value of the formula* Oα at w is the pair [[Oα]]_w = (t, f) defined above. Its support is universal (all ideal alternatives support α) and its opposition is existential (some ideal alternative opposes α): by the clause for P, the opposition of Oα equals the support of P¬α. The *normative status* of α at w is the pair

status_w(α) = (t_{Oα}(w) , t_{O¬α}(w)),

the degree to which α is obligatory and the degree to which α is forbidden. The two pairs differ in general. Take a crisp serial frame in which w has two successors u and v, each with a self-loop, and let α be T at u and F at v. Then Oα = (0,1) and O¬α = (0,1), while Pα = (1,0) and P¬α = (1,0): neither α nor ¬α is obligatory, yet the value of Oα is F. If instead α is B at u and T at v, then Oα = B but O¬α = F.[24] In the general framework, therefore, conflicts between obligations are read from the status pair (both coordinates high), and not from the pair of Oα.

Proposition 0 (Single ideal alternative). If R(w,v*) = 1 for exactly one world v* and R(w,v) = 0 for the others, then status_w(α) = (t_α(v*), f_α(v*)*) and [[Oα]]_w = (t_α(v*), f_α(v*)*) as well.

Proof. With one successor, the support of Oα is t_α(v*) (other terms are* **S***(1, ·) = 1), its opposition is f_α(v*) (other terms are **T**(0, ·) = 0), and the support of O¬α is f_α(v*) by the same argument applied to ¬α. ∎

[22]Kripke (1963), cited in Section 1, for the crisp case. The clauses above are checked against the classical ones by Proposition 1 and by the numerical tests in Section 7.1.
[23]Dunn (1976), cited in Section 1: first-degree entailment is characterised by preservation of the truth order over the four values. The real-valued relation ⩽ generalises it in the obvious way.
[24]Both cases were evaluated with the supplied implementation and agree with the hand calculation. They show that a glut in the value of Oα is not, in general, a conflict between the obligations Oα and O¬α.

The learning study of Section 7 uses exactly this case, so there the pair of an atom at the ideal world, the pair of Oα, and the status of α coincide, and the four corners read: T obligatory, F forbidden, N silent, B conflict.

4. Properties

We state the properties and give proofs. Throughout, 'model' means a model as defined in Section 3.2, formulas may contain free variables under a fixed assignment, and inequalities between values hold coordinatewise as indicated. Numerical checks are reported in Section 7.1; they are sanity checks, not substitutes for the proofs.

Proposition 1 (Conservativity). Let M be a model in which R takes values in {0,1} and every atom takes a value in {T, F} at every world. Then every formula takes a value in {T, F} at every world, and $[\![\varphi]\!]_w$ = T exactly when φ is true at w in the classical Kripke model (W, R, V′) with V′ the corresponding Boolean valuation. Without seriality this is the ordinary semantics of the modal logic K read deontically; with seriality of R it is the semantics of standard deontic logic (KD).

Proof. On {0,1} the operators **T** and **S** coincide with Boolean conjunction and disjunction, and min and max over {0,1} are the Boolean quantifiers. Negation swaps the coordinates and preserves the property t = 1−f. For Oφ the support is $\min_v$ **S**(1−R(w,v), $t_\varphi(v)$); if R(w,v) = 0 the term is **S**(1, t) = 1 and if R(w,v) = 1 it is $t_\varphi(v)$. Hence the support is 1 iff $t_\varphi(v) = 1$ for every v with R(w,v) = 1, which is the classical clause (with vacuous truth when there are no successors). For the opposition, **T**(R, f) is 0 when R = 0 and $f_\varphi(v)$ when R = 1, so the opposition is 1 iff some successor has $f_\varphi(v) = 1$, that is $t_\varphi(v) = 0$, which is the negation of the classical clause. The two coordinates therefore satisfy t = 1−f. ∎ (The proposition is a statement about the evaluation of formulas in models. It does not say that the consequence relation ⊨ of Section 3.2 coincides with classical consequence.)

Proposition 2 (Duality and involution). For all formulas, ¬¬φ and φ have the same value; ¬Oφ and P¬φ have the same value; ¬Pφ and O¬φ have the same value; the analogous De Morgan identities hold for ∧/∨ and ∀/∃.

Proof. Negation swaps coordinates, so applying it twice is the identity. P is defined as ¬O¬, so ¬Oφ = ¬O¬¬φ = P¬φ by involution, and ¬Pφ = ¬¬O¬φ = O¬φ. (Since P is defined as ¬O¬, the second and third identities follow at once from the first.) For ∧/∨, the swapped value of φ∧ψ is (**S**(f_φ,f_ψ), **T**(t_φ,t_ψ)); this is exactly the value of ¬φ∨¬ψ, whose support

is **S**(f_φ,f_ψ) and whose opposition is **T**(t_φ,t_ψ). The argument for $\forall / \exists$ is the same with min and max. ∎

Proposition 3 (Modal monotonicity). If $\varphi \leqslant \psi$ then $O\varphi \leqslant O\psi$ and $P\varphi \leqslant P\psi$.

Proof. Fix a model, a world w and an assignment. By hypothesis $t_\varphi(v) \leq t_\psi(v)$ and $f_\psi(v) \leq f_\varphi(v)$ at every v. Since **S** is non-decreasing in its second argument,[25] $\mathbf{S}(1-R, t_\varphi) \leq \mathbf{S}(1-R, t_\psi)$ for every v, hence the minimum over v is ordered in the same way and the support of Oφ is at most the support of Oψ. Since **T** is non-decreasing in its second argument, $\mathbf{T}(R, f_\psi) \leq \mathbf{T}(R, f_\varphi)$, so the opposition of Oψ is at most that of Oφ. For P, negation reverses $\leqslant$, so $\neg\psi \leqslant \neg\varphi$; applying the first part gives $O\neg\psi \leqslant O\neg\varphi$, and reversing again gives $\neg O\neg\varphi \leqslant \neg O\neg\psi$, that is $P\varphi \leqslant P\psi$. ∎

Proposition 3 is the principle that lets a learner transfer a norm from an action to its consequences: if $\alpha \leqslant \beta$ holds at every ideal world, then $O\alpha \leqslant O\beta$. Two distinct atoms a and b do not satisfy $a \leqslant b$ in all models; the relation must be imposed as background knowledge, and in Section 5 it is imposed as a constraint on sampled contexts. Enforcing sampled inequalities during training does not prove the semantic relation for all contexts.

Proposition 4 (Non-explosion). (a) $p, \neg p \not\models q$. (b) $Op, O\neg p \not\models Oq$, even on models in which every world has at least one successor with R = 1.[26]

Proof. (a) Take a one-world model with V(p) = B and V(q) = F. Then p = (1,1) and ¬p = (1,1) are designated and q = (0,1) is not. (b) Take two worlds w_0, w_1 with $R(w_0,w_1) = R(w_1,w_1) = 1$ and $R(\cdot,w_0) = 0$, so that every world has a successor. Let p = B and q = F at w_1 (values at w_0 are immaterial). At w_0: O p has support $\min(\mathbf{S}(1,\cdot), \mathbf{S}(0,1)) = 1$ and opposition $\mathbf{T}(1,1) = 1$, so Op = B; O¬p has the same value because ¬p = B; but O q has support $\mathbf{S}(0,0) = 0$, so Oq = (0,1) is not designated. ∎

Part (b) matters because in classical Kripke semantics seriality is the frame condition of the D axiom. Here seriality does not cause explosion, because inconsistency lives in the valuation at the ideal world and not in the absence of ideal worlds. *Remark.* The material formula $O\varphi \rightarrow P\varphi$ is not valid in the designated-value sense even on serial frames: on a one-world reflexive frame with φ = N, both Oφ and Pφ are N and the formula is N, which is not designated. Seriality alone therefore does not validate the material D formula, and the consistency penalty

[25] The Łukasiewicz t-norm **T** and t-conorm **S** are continuous and non-decreasing in each argument, and $\mathbf{S}(a,b) = 1 - \mathbf{T}(1-a, 1-b)$ (De Morgan duality under the standard negation). See Hájek (1998), cited in Section 2.2.

[26] In the crisp Kripke semantics, axiom D ($O\varphi \rightarrow P\varphi$) corresponds to seriality of R, a standard frame-correspondence result; this is why classical deontic logics with D cannot tolerate contradictory obligations, and why part (b) of the proposition is stated for serial frames.

used as a baseline in Section 7 is a regulariser inspired by the classical incompatibility of Oφ and O¬φ, and is not the satisfaction loss of that formula.

Proposition 5 (Disjunctive syllogism fails). p, ¬p $\vee$ q $\not\models$ q. Hence the material conditional ¬φ $\vee$ ψ does not support modus ponens in this semantics.

Proof. The model of Proposition 4(a) again: ¬p $\vee$ q = (**S**(1,0), **T**(1,1)) = (1,1), which is designated, p is designated and q is not. ∎

Proposition 5 is a limitation of this conditional in the presence of gluts (it is not a cost that every paraconsistent logic must pay), and it is the reason the specification and training rules in this paper use $\leqslant$ (which is not affected) rather than the material conditional whenever inference is required. In the terminology of the logic of formal inconsistency, a conflict is a *glut* at p, and the failure of disjunctive syllogism is what prevents the glut from spreading to q.[27]

Proposition 6 (Barcan formulas). For a constant finite non-empty domain (the same set $\mathbb{D}$ at every world),[28] the values of O$\forall$xφ and $\forall$xOφ coincide, and so do those of P$\exists$xφ and $\exists$xPφ. (Consequently O$\forall$xφ $\leqslant$ $\forall$xOφ and $\forall$xOφ $\leqslant$ O$\forall$xφ: the Barcan formula and its converse.)

Proof. For the supports: min_v **S**(1−R(w,v), min_d t_φ(v,d)) = min_d min_v **S**(1−R(w,v), t_φ(v,d)), because a non-decreasing function commutes with a finite minimum, and minima commute. For the oppositions: max_v **T**(R(w,v), max_d f_φ(v,d)) = max_d max_v **T**(R(w,v), f_φ(v,d)), by the same argument for **T** and maxima (**T**(a,·) is non-decreasing). The statement for P and $\exists$ follows from Proposition 2. ∎

Proposition 7 (Bounded sensitivity). Fix R. If φ has n atom occurrences, [[φ]] is n-Lipschitz in the sup-norm of the atom valuation. If φ contains k occurrences of O or P, the value is (n+k)-Lipschitz jointly in the atom valuation and R. The map is piecewise linear, so it is differentiable almost everywhere, and at points of differentiability its derivative is bounded in the corresponding operator norm by n+k.[29]

Proof. By induction on φ. Atoms are 1-Lipschitz in their own value. Negation swaps coordinates and is an isometry. For a binary connective, $|\mathbf{T}(a,b) - \mathbf{T}(a',b')| \leq |a-a'| + |b-b'|$,

[27] On gluts and gaps, see Priest (1979) and Carnielli et al. (2007), cited in Section 2.2.
[28] Constant-domain semantics is the setting in which the Barcan formula is valid in classical quantified modal logic; with varying domains it fails. Here the domain is finite and shared by all worlds, so the identities hold exactly.
[29] A function h is L-Lipschitz in the sup-norm if $|h(x) - h(y)| \leq L \cdot \max_i |x_i - y_i|$ for all x, y. The constants in Proposition 7 are upper bounds, not estimates of typical sensitivity.

and the same for **S**, so constants add. Quantifiers and the minima and maxima over worlds are 1-Lipschitz in the sup-norm, so they preserve the constant. For O, the term **S**(1−R, t) is 1-Lipschitz in t and 1-Lipschitz in R separately, which adds at most 1 to the constant per occurrence of a modal operator. All the ingredients are piecewise linear, hence differentiable almost everywhere. ∎

The bound is tight: for a single occurrence of O over a single atom (n = 1, k = 1) the constant is 2 and Section 7.1 finds a ratio of 1.9998 on random perturbations. The proposition concerns the logic layer only. It does not bound the Jacobian of the neural network that produces the atom values, the derivative of the loss, or the total parameter gradient.

Proposition 8 (Complexity). Let φ be a formula of size |φ| with at most q distinct variables, free or bound. All values of φ at all worlds, for all assignments of its variables, can be computed in time O(|φ|·|W|²·|𝔻|^q). For a closed formula, or for a fixed assignment of the free variables, only the assignments to bound variables need be enumerated.[30]

Proof. Each subformula is evaluated for each assignment of its free variables, at most |𝔻|^q assignments. A modal operator at all worlds takes |W|² operations, connectives and quantifiers take O(|W|·|𝔻|) per assignment. Summing over subformulas gives the bound. The memory needed to store all values is O(|φ|·|W|·|𝔻|^q). ∎

Theorem 9 (Containment of inconsistency). Let c(x) = min(t,f) be the *inconsistency degree* of a value x = (t,f).[31] Suppose R takes values in {0,1} and 𝔻 is non-empty. If c([[φ]]_w) > 0 then there are a world v reachable from w by at most md(φ) steps of R (zero steps included, so v = w is allowed), where md is the modal depth, and an atom occurrence of φ, under some assignment of its variables, with c > 0 at v.

Proof. We first show three facts. (1) c(¬x) = c(x). (2) If c(x) = c(y) = 0 then c(x∧y) = 0: if t_x = 0 or t_y = 0 then the support of x∧y is 0; otherwise both supports are positive, so f_x = f_y = 0, so the opposition **S**(f_x,f_y) = 0. By (1) and Proposition 2 the same holds for ∨. (3) For ∀ and ∃, if every instance has c = 0 then c = 0: for ∀, either some instance has t = 0, so min t = 0, or all have t > 0 and therefore f = 0, so max f = 0; ∃ is dual. For O with crisp R, the support is the minimum of t_φ(v) over successors (1 if there are none) and the opposition is the maximum of f_φ(v) over successors (0 if none). If both are positive then every successor has t > 0 and some successor v *has f > 0, so c(*[[φ]]_{v}) > 0. P is dual to O. Now

[30]The bound counts arithmetic operations on real numbers and is for the naive evaluation strategy; it makes no claim about the best possible algorithm.

[31]For the four corners, c is 1 at B and 0 at T, F and N; in the interior it measures how much a value is both supported and opposed. It is not the only reasonable measure of inconsistency, but it is the one for which the closure facts in the proof hold.

the theorem follows by induction on φ: if the value of φ at w is inconsistent then, by (2), (3) or the O-case, some immediate subformula at w or at a successor is inconsistent, and after finitely many steps we reach an atom. ▮

Theorem 9 is the logical basis for the auditing behaviour described in Section 5.3: when a query returns a value with $c > 0$, one can search the derivation backwards for the atoms responsible. It also explains what soft accessibility changes. If R has intermediate values, a single consistent successor can produce an inconsistent O-value: with $R(w,v) = 0.75$ and $V(v) = (0, 0.75)$, we obtain $[[O\varphi]]_w = (0.25, 0.5)$, so $c = 0.25$ although $c = 0$ at v.[32] The soft-R case therefore represents *partial* conflict arising from partial accessibility, and Theorem 9 is a statement about the crisp fragment used in the worked examples and, through the single crisp ideal successor, in the experiments. Two further limits apply. The theorem is an existence statement: it gives a witness atom, not a minimal explanation and not the identity of the source that asserted it. And it concerns exact positivity of c; a sigmoid output is strictly positive, so in a trained network a threshold must be chosen, and the theorem says nothing about thresholded values.

Proposition 10 (A single degree and a conflict). Consider a scalar encoding that assigns to an atom in a context a single truth degree $p \in (0,1)$ and is trained by cross-entropy on $m > 0$ assertions, of which a fraction $\pi \in (0,1)$ assert the atom and the rest deny it. The expected loss is minimised uniquely at $p = \pi$. In particular, one source asserting and one denying ($\pi = 1/2$) yield $p = 1/2$, which is also the degree a scalar model would assign under a uniform prior to a context about which the sources say nothing. A pair-valued semantic model with separate support and opposition targets has its zero-loss optimum at B; with finite sigmoid logits in the supplied network, this is approached as a limit rather than attained exactly.

Proof. The expected loss is $\ell(p) = -\pi \log p - (1-\pi) \log(1-p)$, with derivative $\ell'(p) = -\pi/p + (1-\pi)/(1-p)$, which is negative for $p < \pi$ and positive for $p > \pi$; hence the unique minimiser is $p = \pi$. For the pair-valued model, the support target is 1 (at least one assertion) and the opposition target is 1 (at least one denial), and the two cross-entropy terms are minimised at $t = f = 1$, where both vanish. ▮

Proposition 10 is an example about a specified scalar representation and loss, not a general impossibility result for all scalar encodings or uncertainty-aware learning methods. In that representation there is no state of its own for 'the sources disagree', so a downstream user

[32] Computation: the support is $\mathbf{S}(1-0.75, 0) = 0.25$ and the opposition is $\mathbf{T}(0.75, 0.75) = \max(0, 0.75+0.75-1) = 0.5$.

cannot tell disagreement from indifference, while the pair-valued model has such a state. Table 2 summarises the results of this section.[33]

Result	Statement	Assumption on R	Used in
Prop. 0	Single ideal alternative: status equals pair	one crisp successor	Section 7
Prop. 1	Agreement with classical Kripke evaluation	crisp	Section 7.1
Prop. 2	Involution and modal duality	any	Section 3
Prop. 3	Modal monotonicity under truth-order entailment $\preccurlyeq$	any	Section 5.2
Prop. 4	Non-explosion, also on serial frames	crisp example	Section 6
Prop. 5	Disjunctive syllogism fails	–	Section 8
Prop. 6	Barcan formula and its converse	any	Section 3
Prop. 7	Bounded sensitivity (n+k)	any	Section 5.1
Prop. 8	Time $O(\|\varphi\|\cdot\|W\|^2\cdot\|\mathbb{D}\|^q)$	any	Section 7.2
Thm. 9	Inconsistency traceable to atoms	crisp	Section 5.3
Prop. 10	Scalar encoding under cross-entropy: $p = \pi$	–	Section 7.3

Table 2. Results of Section 4 and the assumption on R used in each proof.

5. Learning with PDL-$\mathbb{R}$

5.1 Architecture

The neural component is a function g_θ that maps an input context x (a feature vector describing the situation in which norms apply) to the values of atoms at the worlds of the model. In the general architecture, each world $v \in W$ has an embedding e_v, the valuation at v is produced by heads on top of $g_\theta(x)$ and e_v, and accessibility is $R_\vartheta(w,v) = \sigma(e_w^T \Theta e_v)$ for a learned matrix Θ, where σ is the logistic function.[34] Each atom at each world has two heads, one for support and one for opposition, each followed by a sigmoid, so that the pair lies in $[0,1]^2$. Everything above the heads is the fixed semantics of Section 3, so formulas can be evaluated on the network output and differentiated end to end. Proposition 8 gives the cost of the logical part and Proposition 7 bounds the sensitivity of the logic layer; neither bounds the gradient of the neural component. The experiments of Section 7 use a reduction of this architecture, described there, and do not exercise learned accessibility or quantifiers.

5.2 Training signals

[33]The table lists what each property says and the assumption on R that its proof uses; 'any' means R may take any value in [0,1].
[34]This bilinear parameterisation is one choice among many; any differentiable map from pairs of world embeddings to [0,1] can serve, and crisp accessibility can be obtained by thresholding or by fixing R. The experiments in Section 7 fix a two-world frame with R crisp.

Three kinds of signal are used. *Labelled assertions* (full-status supervision): for each observed cell, the support head receives target 1 if some source asserts the formula and 0 otherwise, and the opposition head receives target 1 if some source denies it and 0 otherwise, by a cross-entropy term,[35] Two sources may assert φ and deny φ in the same context; since support and opposition are separate coordinates, both targets can be fitted and the value of φ tends to B. In this supervision, silence in a fully observed cell is treated as a target of 0 on both heads; cells that are not observed contribute no term. *Entailment constraints*: for a known rule $\varphi \leqslant \psi$, the violation on a context is $\nu(\varphi,\psi) = \max(0, t_\varphi - t_\psi) + \max(0, f_\psi - f_\varphi)$, and its mean is added to the loss. By Proposition 3 the truth-order relation is preserved by O, so a constraint imposed at the level of the atoms at the ideal world propagates to the obligations. *Consistency regularisation* (optional): the quantity $\mathrm{relu}(t + f - 1)$ measures how far a value is from having $t + f \leq 1$. Penalising it discourages gluts, in the spirit of the classical incompatibility of Oφ and O¬φ, but it is not the satisfaction loss of the material D formula (Section 4, remark after Proposition 4); penalising $c(x) = \min(t,f)$ directly is a softer alternative. The experiments below use the first two signals for the configuration P, and all three for the configuration C, which is a consistency-regularised baseline and not a general classical-logic competitor.

5.3 A proposed interface for auditing

Theorem 9 suggests an interface, which we describe as a proposal: it is not implemented in the supplied code and it is not evaluated. For crisp accessibility, when a query about a context returns a value with $c > 0$, the derivation can be walked backwards through the formula, and at each step the subformula and world that carry inconsistency are identified until an atom with $c > 0$ is reached. In a trained network c is strictly positive everywhere, so a decision rule is needed: in the experiments the score $\min(t,f)$ is used for detection and its AUC is measured, and a deployed system would need a threshold and a calibration policy chosen on validation data. What a trace certifies is the existence of a witness atom, not the identity of the source that asserted it, which would have to be recorded in the data. A conflict found this way is a case for a person with authority to prioritise norms.[36]

5.4 What the logic does not do

PDL-$\mathbb{R}$ does not resolve conflicts. It has no preference or priority relation among norms, and Section 8 says what that means for use. It also does not internalise a consistency operator in

[35]For a target $y \in \{0,1\}$ and a predicted degree p, the cross-entropy is $-y \log p - (1-y) \log(1-p)$; we clamp p to $[10^{-5}, 1-10^{-5}]$ for numerical stability.

[36]The traceability is a property of the semantics (Theorem 9) for crisp accessibility and not of any particular network. Source provenance, minimal explanations and thresholded guarantees would need additional definitions and implementation.

the object language, as Logics of Formal Inconsistency do, and it inherits the failure of disjunctive syllogism proved in Proposition 5.

6. Two worked examples

The values below were computed by executing the semantics of Section 3 on the models described, and match hand calculation.

6.1 Conflicting obligations from two sources

Let r be the proposition 'the record is retained'. Source A requires retention and source B requires deletion. Take $W = \{w_0, w_1\}$ with $R(w_0,w_1) = R(w_1,w_1) = 1$ and $R(\cdot,w_0) = 0$, so that w_1 is the single ideal alternative, and let r = B and q = F at w_1, where q is an unrelated proposition. Then:

Formula at w_0	Support	Opposition	Value
O r	1	1	B (conflict, designated)
O ¬r	1	1	B (conflict, designated)
O q	0	1	F (not designated)

Both obligations are derived and reported as conflicting, and the unrelated q is not derived, as Proposition 4(b) states. In classical SDL both r and ¬r would be obligatory only in an inconsistent theory, and every q would be obligatory as well.[37]

6.2 Chisholm's contrary-to-duty set

Let g be 'Jones goes' and s be 'Jones tells'. Chisholm's four sentences are: (1) O g; (2) O(g → s); (3) ¬g → O¬s; (4) ¬g.[38] In SDL, (1) and (2) give O s by closure under implication, and (3) with (4) gives O¬s by detachment, which is where the difficulty arises. We take the two-world frame of Section 6.1, read → as $\neg\varphi \vee \psi$, and examine two models. In model M_1, g = T and s = B at the ideal world w_1, and g = F, s = F at w_0. In model M_2, g = B and s = F at w_1, with the same values at w_0. Values at w_0:

Formula at w_0	Model M_1	Model M_2	Designated in both?
(1) O g	T	B	yes
(2) O(g → s)	B	B	yes
(3) ¬g → O¬s (= g ∨ O¬s)	B	T	yes
(4) ¬g	T	T	yes
O s	B	F	no
O q, q unrelated (F at w_1)	F	F	no

[37] This is the standard explosion of SDL: with axiom D, Oφ → ¬O¬φ, the set {Or, O¬r} is inconsistent, and in classical logic an inconsistent set entails everything. See von Wright (1951), cited in Section 1.
[38] Chisholm (1963), cited in Section 2.3.

Table 3. Chisholm's set in two models of PDL-$\mathbb{R}$ (all values computed with the supplied implementation).

All four premises are designated in both models, so the set is satisfiable. In M_1 the obligations O s and O¬s are both designated, with the conflict located at the atom s at the ideal world; in M_2 the premises hold and O s is not designated. The premises therefore do not entail O s, and they do not fix where an inconsistency lies. In neither model is the unrelated obligation O q designated, as Proposition 4 requires. The example shows that Chisholm's set is satisfiable without collapse; it does not show that a learner given only these premises would place a conflict at s, and it does not dissolve the paradox: the standard responses, such as dyadic deontic logics in which (3) is a conditional obligation whose antecedent is the violation, use a structure that PDL-$\mathbb{R}$ does not have. If the derivation of O s were wanted, explicit order constraints or consistency assumptions would have to be added and the resulting consequence proved.

7. Experiments

The semantics of Section 3 is implemented in PyTorch, and the supplementary package contains property checks, a scaling experiment, worked examples, the synthetic learning script and per-seed summary metrics. The revised statistical analysis is computed from those archived metrics; the 300 learning runs were not repeated for this revision.

7.1 Numerical checks of the properties

Each property of Section 4 was checked on randomly generated models: worlds between 2 and 8, domains between 2 and 5 elements, and accessibility and valuations drawn uniformly from [0,1] unless a crisp model was required. Table 4 gives the results. No violation was found. In addition, 20 000 random nested formulas (depth up to 3, connectives ¬, $\wedge$, $\vee$, O, P over three atoms taking the four corner values, crisp random frames with 2 to 5 worlds) were used to check Proposition 2, Proposition 3 and Theorem 9. A second test used 5 000 nested formulas with continuous-valued, individually consistent atoms and crisp frames; no formula became inconsistent, checking the contrapositive of Theorem 9 beyond the four corners.

Property	Trials	Result
Prop. 1: agreement with classical Kripke semantics (O and P, crisp models)	4000	0 disagreements
Prop. 2: ¬Oφ and P¬φ identical	2000	max difference 0
Prop. 3: $\varphi \leqslant \psi$ implies $O\varphi \leqslant O\psi$	3000	0 violations

Property	Trials	Result
Prop. 6: O∀xφ and ∀xOφ identical	2000	max difference 0
Prop. 7: ratio of output change to input change (bound 2 for n=1, k=1)	3000	max ratio 1.9998
Prop. 4(b): countermodel (Section 4)	1	Op = B, O¬p = B, Oq = F
Nested formulas: Prop. 2 and Prop. 3 (O and P over φ ⩽ φ∨χ)	20000	0 + 0 violations
Nested formulas: Thm. 9 (crisp R; 12961 inconsistent cells)	20000	0 violations

Table 4. Numerical checks of the properties. Random tests cannot prove a property, but a violation would have refuted an implementation or a proof.[39]

7.2 Scaling

Proposition 8 predicts quadratic growth of the number of operations for evaluating O over |W| worlds. We timed the evaluation of Oφ at all worlds for |W| from 64 to 4096 on one CPU thread (median of 15 timed repetitions after 3 warm-ups, PyTorch 2.14.0+cu130). The fitted slope of log-time against log-size is 2.20 over all sizes and 2.14 over the four largest sizes. Median times were 64: 0.04 ms; 128: 0.07 ms; 256: 0.24 ms; 512: 2.65 ms; 1024: 10.16 ms; 2048: 46.48 ms; 4096: 225.14 ms. The fitted timing slope depends on the machine and runtime and is not itself a proof of quadratic complexity. Proposition 8 establishes the operation count for the stated naive implementation; these timings are illustrative only.

7.3 Learning with conflicting normative sources

Task. Contexts $x \in \mathbb{R}^{10}$ are drawn from a standard normal distribution. There are m = 5 action pairs (a_j, b_j). A hidden linear score determines the status of a_j in each context: obligatory if the score exceeds 0.6, forbidden if it is below −0.6, and otherwise silent. The action b_j is a necessary consequence of a_j, which we write a_j ⩽ b_j, so that O a_j ⩽ O b_j and O¬b_j ⩽ O¬a_j; b_j has its own hidden score that decides its status when a_j is silent. Two normative sources overlap in a region of the input space defined by a third hidden score: in that region source 1 asserts O a_j and source 2 asserts O¬a_j, so the true value of a_j is B. Each context has five such cells, and the fraction of cells that fall in the conflict region is varied from 0 to 0.33 by moving the threshold; the resulting fractions are the values reported on the horizontal axis.[40] In the conflict region b_j is obligatory, since O a_j holds and O a_j ⩽ O b_j.

[39]The proofs are in Section 4. The tests exist to catch algebra errors in the proofs or in the code, and each was run with a fixed random seed that is set in the supplied script.

[40]The threshold is the normal quantile of 1 − ρ for a nominal rate ρ, but the hidden score has a standard deviation above one, so realised rates exceed the nominal ones (0.05, 0.10, 0.20, 0.30 give 0.090, 0.144, 0.238, 0.327). We report the realised rates throughout.

Data and supervision. The training set has 300 contexts with all a-labels observed, including both assertions in conflict cells. Labels for b are observed for a random 10% of non-conflict cells only, so most of the knowledge about b must come from the entailment rule. In addition, 1500 unlabelled contexts are used for the rule loss. The test set has 3000 contexts sampled from the same distribution, including conflict regions, with true statuses in {N, T, F, B} for both a and b.

Model. A multilayer perceptron with two hidden layers of 64 tanh units maps x to twenty sigmoid outputs: for each of the ten atoms a_j, b_j, a support head and an opposition head. In the notation of Section 5 the network gives the value of each action at a single ideal world, so that O a_j is the pair (support of a_j, opposition of a_j) at that world, and the status is read off by thresholding both coordinates at 0.5. Training uses Adam[41] with learning rate 0.003 for 500 full-batch epochs. All three models share the architecture and the hyperparameters, and there was no tuning of any of them.[42]

Compared configurations. N (data only): cross-entropy on the observed labels, no logic. C (rules plus consistency regulariser): data, entailment constraints, and the penalty relu(t + f − 1) on every atom, a regulariser inspired by the classical incompatibility of O a and O¬a (Section 4); it is a baseline of that specific kind and not a general classical-logic competitor. P (rules, no consistency regulariser): data and entailment constraints. Configurations C and P differ in exactly one term of the loss, and P is exactly a two-head classifier with rule regularisation. The loss terms and their normalisation are given in Appendix A.

Metrics. Four-way status accuracy on a and on b over all test cells; the area under the ROC curve for detecting conflict cells, using the inconsistency degree min(t,f) of a as the score; results are means and standard deviations over 20 seeds, where a seed determines the hidden scores, the data and the initialisation. Paired differences between P and each baseline were tested with the exact two-sided Wilcoxon signed-rank test (full enumeration of sign flips, average ranks for ties, zero differences dropped),[43] and are reported with 95% t-intervals over the 20 seeds. The p-values are unadjusted and exploratory; the family of comparisons is the accuracy and AUC comparisons of P against C and N at each conflict level.

Results. Table 5 and Figure 1 give the results.

---

[41]Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*. arXiv:1412.6980.

[42]Fixing the hyperparameters in advance for all models avoids giving any configuration an advantage from tuning, at the cost of not knowing how far each could be improved.

[43]Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83. The exact test is implemented in the supplied script; accuracies are counts out of 15 000 cells, so differences are exact multiples of 1/15 000.

Conflict rate	a: N	a: C	a: P	b: N	b: C	b: P
0.000	0.887 ± 0.007	0.889 ± 0.005	0.888 ± 0.006	0.709 ± 0.013	0.731 ± 0.013	0.727 ± 0.013
0.090	0.798 ± 0.012	0.785 ± 0.016	0.800 ± 0.014	0.665 ± 0.016	0.706 ± 0.010	0.699 ± 0.009
0.144	0.779 ± 0.010	0.747 ± 0.013	0.780 ± 0.011	0.643 ± 0.015	0.702 ± 0.009	0.691 ± 0.010
0.238	0.771 ± 0.012	0.687 ± 0.015	0.771 ± 0.011	0.608 ± 0.017	0.706 ± 0.009	0.690 ± 0.017
0.327	0.777 ± 0.009	0.633 ± 0.015	0.777 ± 0.008	0.570 ± 0.026	0.720 ± 0.012	0.690 ± 0.019

Table 5. Four-way status accuracy (mean ± standard deviation over 20 seeds) for the source-level norm a and the derived norm b. N: data only; C: rules and consistency penalty; P: rules, paraconsistent. The conflict rate is the mean fraction of test cells that lie in the conflict region.

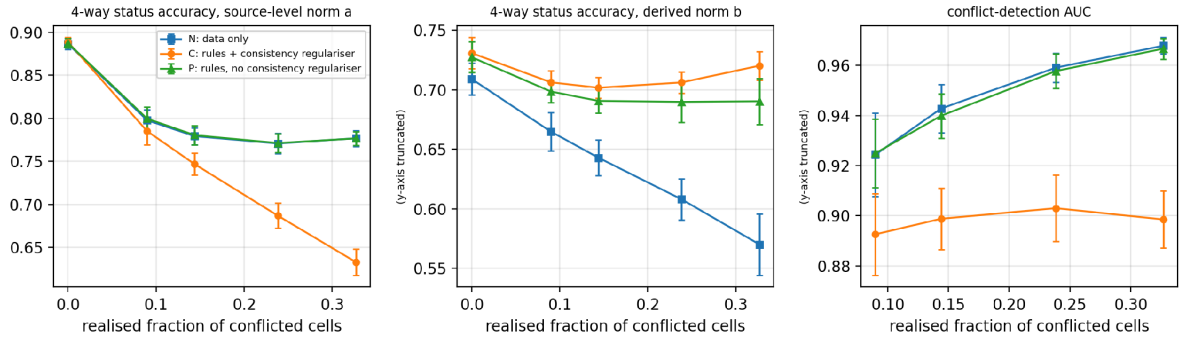


Figure 1. Results of the learning study. Left: status accuracy for a. Middle: status accuracy for b. Right: AUC for detecting conflict cells. Error bars are one standard deviation over 20 seeds; the seed varies both the generated task and the initialisation. The vertical axes are truncated and do not start at zero.[44]

Conflict rate	AUC: N	AUC: C	AUC: P
0.090	0.924 ± 0.017	0.893 ± 0.016	0.925 ± 0.014
0.144	0.943 ± 0.010	0.899 ± 0.012	0.940 ± 0.009
0.238	0.959 ± 0.006	0.903 ± 0.013	0.958 ± 0.007
0.327	0.968 ± 0.003	0.898 ± 0.011	0.967 ± 0.004

Table 6. Conflict-detection AUC (mean ± standard deviation over 20 seeds).

Observations. First, with no conflicts (rate 0), the three configurations are close on a (0.887 to 0.889), and the two configurations that use the entailment rule are more accurate on b (0.727 for P and 0.731 for C) than the data-only model (0.709); the P–N difference is +0.018 (all 20 seeds, exact $p < 0.001$). Second, the consistency-regularised configuration C loses accuracy on a as conflicts increase: at a realised conflict rate of 0.327 it is at 0.633 against

[44]The figure is generated by the supplied script from the raw per-seed results file.

0.777 for P, and P is ahead of C in all 20 seeds at rates of 0.144 and above (+0.144 at 0.327, 95% interval 0.136 to 0.151). Third, the data-only model N has the same accuracy on a as P (differences between −0.0001 and +0.0019, none significant), but its accuracy on the derived norm b falls from 0.709 to 0.570 as conflicts increase, while P stays near 0.69; at 0.327 the gain of P over N is +0.120 (interval 0.109 to 0.132) in all 20 seeds. Fourth, C is better than P on the derived norm b at every tested conflict level: the mean gap (C − P) is 0.003 at rate 0 and grows to 0.030 at 0.327, and P is below C in 16 to 19 of the 20 seeds at each level. The comparison is therefore a trade-off and not a dominance: P is better on the source-level norm a and on conflict detection, and C is better on b. Fifth, on conflict detection, the mean AUCs of N and P are similar (0.924 to 0.968), with paired differences P − N of +0.0006, −0.0029, −0.0014 and −0.0012 at the four conflict levels; only the second (interval −0.0048 to −0.0010, unadjusted exact p = 0.006) is distinguishable from zero, and it is small. C has a lower AUC than P at every level, by 0.032 to 0.068 (20 of 20 seeds at each level). We do not claim equivalence between N and P on AUC; the study has no pre-specified margin.

Interpretation. The pattern is what the logic suggests, and the study is consistent with the following reading, which it does not prove by ablation.[45] Both P and N can represent B, so both fit the conflicting assertions on a and can flag them. Only the configurations with rules propagate norms to b from the few labels that exist. C has the rules, but its consistency regulariser is violated by every conflict cell, which coincides with lower accuracy on a and lower detection AUC. C is nevertheless better than P on b, where every cell is conflict-free by construction; we did not isolate why, and one reading is that the regulariser acts on the b heads as well, but this was not tested.

7.4 Scope of the experiments

The study evaluates a reduction, and its scope should be read narrowly.[46] The training script does not call the semantic implementation of Section 3 and contains no modal layer, no world embeddings, no learned accessibility and no quantifiers: it trains a two-head multilabel network with two inequality constraints (and, for C, a consistency regulariser), which is what the single-crisp-ideal-successor case reduces to (Proposition 0). The study therefore says nothing about the practical value of multi-world or first-order reasoning, learned accessibility, source tracing, or Theorem 9 on a non-trivial formula graph. The derived norm b is generated

[45]The prediction is qualitative: the logic says which configurations can represent B and which impose rules, not by how much accuracy differs. The reading offered here would need further ablations, for example a sweep of the consistency weight chosen on validation data, to be tested.

[46]A synthetic design was chosen because the true value (N, T, F or B) of every cell is then known, which is not the case for naturally occurring normative corpora, where conflicts are themselves a matter of interpretation.

to satisfy the two inequalities exactly, b never has the conflict status, and its labels are withheld in conflict cells, so the evaluation is a test of extrapolation through the rules; that C does better on b is unsurprising for that reason, and the explanation was not tested. The hidden scores are linear and the conflict region is a half-space, so detection is learnable by construction. Hyperparameters were fixed in advance and identical across models, with no sweep of the consistency weight and no validation set; different tuning might change the sizes of the differences. There are 20 seeds and one architecture, and no comparison with logic-tensor-network, logical-neural-network or modal-network implementations, so no claim of superiority over them is made. Accuracy is reported over all cells and not per class or balanced; a detection threshold, precision and recall were not evaluated.

8. Discussion

What the logic gives. The semantics makes four situations distinguishable in the output of a learner: silence, support, opposition, and conflict. For crisp accessibility, Theorem 9 gives an existence witness for an inconsistent atom at a reachable world, without identifying its source; Proposition 4 shows that an obligation conflict does not by itself entail an unrelated obligation. In the study of a single-ideal-world reduction, dropping the consistency regulariser went with better accuracy on the norms that are in conflict, while the rule constraint went with better accuracy on derived norms than data alone; the consistency-regularised model remained better on derived norms.

Knowledge engineering. Writing a normative specification for PDL-$\mathbb{R}$ requires choosing which formulas are atoms at ideal worlds, and stating necessary-consequence rules as entailments ($\leqslant$) instead of material conditionals. The second requirement is a real burden on the modeller: by Proposition 5, a rule written with $\neg\varphi \vee \psi$ does not support modus ponens in the presence of a glut, and a truth-order rule $\varphi \leqslant \psi$ is not valid in all models but has to be imposed. Tooling that checks whether a rule holds in the truth order, on samples or symbolically, would help.[47]

Human in the loop. The semantics can report a conflict without silently resolving it. A future tracing implementation could identify a witness atom under the crisp-frame conditions of Theorem 9, but the current code does not trace source provenance or provide a user-facing

[47] A sample-based check evaluates $\nu(\varphi,\psi)$ of Section 5.2 on a set of contexts or on a grid of atom values and reports the largest violation; this can falsify $\varphi \leqslant \psi$ but cannot prove it. Checking the four corner values of each atom is not sufficient: for $\varphi = p$ and $\psi = p \wedge p$ the two formulas agree at all four corners, yet at $p = (0.5, 0)$ the values are $(0.5, 0)$ and $(0, 0)$, so $p \leqslant p \wedge p$ fails. An exact check would have to encode all min, max and clamp branches, for example as linear or mixed-integer constraints, and its cost is a separate question.

explanation. Choosing priority between normative sources remains a decision for an authorised person. Adding a priority relation would require the guarantees of Section 4 to be checked again.[48]

Scalability. The modal layer is quadratic in the number of worlds (Proposition 8), and quantification is exponential in the number of nested variables over the domain. The learning experiment uses one ideal world, while the property checks and examples use small frames (Section 7.2); we did not study realistic normative specifications, so we make no claim about the number of worlds they need. For large domains, sampling or grounding strategies of the kind used in logic tensor networks would be needed.[49]

Limits of the semantics. Three limitations are inherent to the design. The consequence relation is not closed under modus ponens for the material conditional (Proposition 5). Accessibility is not required to be serial, so a world without ideal successors makes every obligation vacuously true and every permission false, as in classical Kripke semantics; a modeller who wants D as a frame condition can impose seriality on R. Finally, PDL-$\mathbb{R}$ has no temporal, epistemic or defeasible component and no representation of conditional obligation, which is why Chisholm's set (Section 6.2) is treated only as a satisfiability example.

9. Conclusion

We presented a finite-domain, pair-valued deontic semantics and related its propositional and modal clauses to interval-bound formulations. Its formal results establish Boolean Kripke agreement on crisp frames, non-explosion, and an inconsistency-containment guarantee under crisp accessibility. The synthetic learning study evaluates a single-ideal-world reduction: rule constraints improve derived-norm accuracy over data-only training, while omitting a consistency regulariser improves source-norm accuracy and conflict detection at higher conflict rates but reduces derived-norm accuracy relative to the regularised configuration. This trade-off is evidence about the tested reduction, not about an implemented multi-world or first-order neural reasoner. The next steps are a neural implementation and evaluation of multi-world quantified inference and learned accessibility, source-level provenance for

[48]In particular, a priority relation that discards the lower-ranked norm changes the value of Oφ and therefore Propositions 3 and 4, and Theorem 9 would have to be restated for the modified operator.
[49]Quantifiers over large domains are the usual scalability concern for real-valued first-order logics; the finite-domain min and max of Section 3 can be replaced by smooth aggregators over sampled instances, at the cost of Proposition 6 holding only approximately.

conflicts, comparison with interval-bound networks, and tests on naturally conflicting normative data.[50]

[50]The code, the data generator and the raw results (per seed and configuration) accompany the submission as supplementary material, so that every number in Tables 4 to 6 and in Figure 1 can be reproduced.

Code and data availability

The code, data generators, and per-seed summary metrics are supplied as supplementary material. A version-controlled copy is deposited at https://github.com/sgurfinkel1-dev/paraconsistent-deontic-networks (Godoy, 2026; commit c6d7fb305c979d8e358e5c6f5f92f2d9d7a1d965). Raw generated examples and model checkpoints are not archived.

References

Works cited only in footnotes are also listed here.[51]

Badreddine, S., d'Avila Garcez, A., Serafini, L., & Spranger, M. (2022). Logic tensor networks. *Artificial Intelligence, 303*, Article 103649. https://doi.org/10.1016/j.artint.2021.103649

Belnap, N. D. (1977). A useful four-valued logic. In J. M. Dunn & G. Epstein (Eds.), *Modern uses of multiple-valued logic* (pp. 5–37). Springer Netherlands. https://doi.org/10.1007/978-94-010-1161-7_2

Bílková, M., Frittella, S., & Kozhemiachenko, D. (2022). *Paraconsistent Gödel modal logic*. arXiv:2203.01237. https://arxiv.org/abs/2203.01237

Boscarato, L., Donadello, I., Artale, A., Montali, M., & Maggi, F. M. (2026). *First-order temporal logic tensor networks*. arXiv:2606.29972. https://arxiv.org/abs/2606.29972

Carnielli, W., Coniglio, M. E., & Marcos, J. (2007). Logics of formal inconsistency. In D. M. Gabbay & F. Guenthner (Eds.), *Handbook of philosophical logic* (Vol. 14, pp. 1–93). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6324-4_1

Chisholm, R. M. (1963). Contrary-to-duty imperatives and deontic logic. *Analysis, 24*(2), 33–36.

da Costa, N. C. A. (1974). On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic, 15*(4), 497–510. https://doi.org/10.1305/ndjfl/1093891487

Dunn, J. M. (1976). Intuitive semantics for first-degree entailments and 'coupled trees'. *Philosophical Studies, 29*(3), 149–168.

Garcez, A. d., & Lamb, L. C. (2023). Neurosymbolic AI: The 3rd wave. *Artificial Intelligence Review, 56*(11), 12387–12406. https://doi.org/10.1007/s10462-023-10448-w

Garcez, A. S. d'Avila, Broda, K. B., & Gabbay, D. M. (2002). *Neural-symbolic learning systems: Foundations and applications*. Springer.

Garcez, A. S. d'Avila, Lamb, L. C., & Gabbay, D. M. (2007). Connectionist modal logic: Representing modalities in neural networks. *Theoretical Computer Science, 371*(1–2), 34–53. https://doi.org/10.1016/j.tcs.2006.10.023

Godoy, S. G. M. (2026). Pair-valued deontic semantics: Research code and results [Computer software]. GitHub. https://github.com/sgurfinkel1-dev/paraconsistent-deontic-networks

Goble, L. (2009). Normative conflicts and the logic of 'ought'. *Noûs, 43*(3), 450–489. https://doi.org/10.1111/j.1468-0068.2009.00714.x

Hájek, P. (1998). *Metamathematics of fuzzy logic*. Kluwer Academic Publishers.

Kifer, M., & Subrahmanian, V. S. (1992). Theory of generalized annotated logic programming and its applications. *The Journal of Logic Programming, 12*(4), 335–367.

Kingma, D. P., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations*. arXiv:1412.6980.

Kripke, S. A. (1963). Semantical analysis of modal logic I: Normal modal propositional calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 9*(5–6), 67–96. https://doi.org/10.1002/malq.19630090502

Manhaeve, R., Dumančić, S., Kimmig, A., Demeester, T., & De Raedt, L. (2021). Neural probabilistic logic programming in DeepProbLog. *Artificial Intelligence, 298*, Article 103504. https://doi.org/10.1016/j.artint.2021.103504

[51]Reference entries follow APA style. DOIs were resolved against Crossref while preparing the submission where a DOI exists.

Odintsov, S. P., & Wansing, H. (2010). Modal logics with Belnapian truth values. *Journal of Applied Non-Classical Logics, 20*(3), 279–301. https://doi.org/10.3166/jancl.20.279-301

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., … Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32* (pp. 8024–8035).

Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic, 8*(1), 219–241.

Riegel, R., Gray, A., Luus, F., Khan, N., Makondo, N., Akhalwaya, I. Y., … Srivastava, S. (2020). *Logical neural networks*. arXiv:2006.13155. https://arxiv.org/abs/2006.13155

Sulc, A., & Naddour, N. (2025). *Modal logic neural networks*. arXiv:2512.03491 (version 3). https://arxiv.org/abs/2512.03491

von Wright, G. H. (1951). Deontic logic. *Mind, 60*(237), 1–15.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*(6), 80–83.[52]

[52]Cited for the signed-rank test; the exact-enumeration implementation used in Section 7.3 is in the supplied script stats.py.

Appendix A. Implementation

The semantics is implemented on tensors. A formula value is a pair of tensors (t, f). The Łukasiewicz operators are **T**(a,b) = clamp(a+b−1, min = 0) and **S**(a,b) = clamp(a+b, max = 1). For the modal operator, the value tensors have shape [|W|] and the accessibility tensor has shape [|W|,|W|], so that O is computed with one broadcasted operation and one reduction per coordinate:[53]

```
box(vt, vf, R):
    t = clamp(1 - R + vt[None,:], max=1).min(over worlds v)
    f = clamp(R + vf[None,:] - 1, min=0).max(over worlds v)
    return (t, f)

dia(vt, vf, R):   # P phi = not O not phi
    (bt, bf) = box(vf, vt, R)
    return (bf, bt)

forall(vt, vf, dim): return (vt.min(dim), vf.max(dim))
exists(vt, vf, dim): return (vt.max(dim), vf.min(dim))
```

The loss used in the experiments is the sum of four cross-entropy terms, each averaged over its cells (the support and opposition heads of the a-atoms over all training cells; those of the b-atoms over the labelled b-cells), plus, for C and P, λ_rule times the mean over the training and unlabelled contexts and the five action pairs of relu(t_a − t_b) + relu(f_b − f_a), plus, for C only, λ_D times the sum of two means, one of relu(t + f − 1) over the a-atoms and one over the b-atoms. In the experiments λ_rule = λ_D = 1.[54]

Appendix B. Experimental protocol

Item	Value
Context dimension; action pairs	10; 5
Score thresholds	±0.6 for the status of a; 0.5 for b when a is silent; 0 for the forbidden case of b
Training contexts; unlabelled contexts; test contexts	300; 1500; 3000
Fraction of non-conflict b-cells labelled	10%
Network	10–64–64–20, tanh hidden units, sigmoid outputs
Optimiser	Adam, learning rate 0.003, 500 full-batch epochs
Loss weights	λ_rule = 1 (N: 0); λ_D = 1 (only C)
Seeds	20 per configuration and conflict level (seeds 0 to 19)
Nominal conflict levels	0, 0.05, 0.10, 0.20, 0.30 (realised: 0, 0.090, 0.144, 0.238, 0.327)
Status decision	threshold 0.5 on each coordinate

[53]The listing is a paraphrase of the supplied source file, in the same order of operations; the file is the reference implementation.
[54]The value 1 was chosen in advance and not tuned; see Section 7.3.

Appendix C. Additional results

Table C1 reports accuracy on the source-level norm a outside the conflict region and the mean inconsistency degree min(t,f) assigned to those cells. The latter is an average score of simultaneous support and opposition, not the frequency of reported conflicts; no detection threshold was applied.

Conflict rate	Acc. a, non-conflict cells: N	C	P	Mean inconsistency, non-conflict cells: N	C	P
0.000	0.887 ± 0.007	0.889 ± 0.005	0.888 ± 0.006	0.002 ± 0.000	0.001 ± 0.000	0.001 ± 0.000
0.090	0.830 ± 0.010	0.846 ± 0.010	0.836 ± 0.012	0.022 ± 0.005	0.010 ± 0.002	0.019 ± 0.005
0.144	0.810 ± 0.010	0.834 ± 0.009	0.817 ± 0.011	0.032 ± 0.005	0.013 ± 0.002	0.029 ± 0.005
0.238	0.780 ± 0.012	0.811 ± 0.010	0.787 ± 0.012	0.048 ± 0.006	0.018 ± 0.002	0.044 ± 0.006
0.327	0.756 ± 0.014	0.787 ± 0.009	0.760 ± 0.012	0.061 ± 0.006	0.022 ± 0.002	0.059 ± 0.006

Table C1. Non-conflict cells: accuracy on a, and mean inconsistency degree (mean ± standard deviation over 20 seeds).

The consistency-regularised configuration C assigns a smaller mean inconsistency degree to non-conflict cells (at most 0.022 at the highest conflict rate, against 0.061 for N and 0.059 for P), which is what its penalty is for; the cost is on the conflict cells themselves, as Table 5 shows. The mean of min(t,f) is a mean score and not a frequency of reported conflicts; no threshold was applied to it. Tables C2a, C2b and C3 give the paired differences, 95% intervals and exact Wilcoxon signed-rank p-values behind the statements in Section 7.3.[55]

Rate	P − C on a (95% interval)	p	seeds P>C
0.000	-0.0013 [-0.0026, 0.0001]	0.0651	7/20
0.090	0.0145 [0.0105, 0.0186]	<0.0001	19/20
0.144	0.0338 [0.0296, 0.0379]	<0.0001	20/20
0.238	0.0844 [0.0782, 0.0907]	<0.0001	20/20
0.327	0.1438 [0.1364, 0.1512]	<0.0001	20/20

[55] p-values are two-sided, from full enumeration of the sign-flip distribution of the signed-rank statistic with average ranks for ties, and unadjusted for multiple comparisons. With 20 non-zero pairs the smallest attainable value is $2/2^{20}$, about 1.9×10^{-6}; entries shown as <0.0001 are below the reporting resolution and not necessarily equal to that minimum. Confidence intervals are t-intervals over the 20 seeds.

Table C2a. Paired difference in four-way status accuracy on the source-level norm a (P − C) with 95% intervals, exact Wilcoxon signed-rank p-values and the number of seeds (out of 20) in which P is strictly better. Rate is the realised fraction of conflicted cells.

Rate	P − N on b	p	seeds P>N	P − C on b	p	seeds P>C
0.000	0.0183 [0.0134, 0.0232]	<0.0001	20/20	-0.0035 [-0.0051, -0.0019]	0.0003	4/20
0.090	0.0337 [0.0261, 0.0412]	<0.0001	20/20	-0.0077 [-0.0105, -0.0050]	<0.0001	1/20
0.144	0.0477 [0.0392, 0.0562]	<0.0001	20/20	-0.0110 [-0.0150, -0.0071]	<0.0001	1/20
0.238	0.0819 [0.0725, 0.0913]	<0.0001	20/20	-0.0163 [-0.0226, -0.0100]	<0.0001	4/20
0.327	0.1203 [0.1089, 0.1317]	<0.0001	20/20	-0.0299 [-0.0387, -0.0211]	<0.0001	1/20

Table C2b. The same statistics for the derived norm b, for P − N and P − C.[56]

Conflict rate	P − N on AUC	p	seeds P>N	P − C on AUC	p	seeds P>C
0.090	0.0006 [-0.0022, 0.0033]	0.7012	12/20	0.0324 [0.0285, 0.0363]	<0.0001	20/20
0.144	-0.0029 [-0.0048, -0.0010]	0.0056	5/20	0.0411 [0.0380, 0.0442]	<0.0001	20/20
0.238	-0.0014 [-0.0031, 0.0004]	0.1536	8/20	0.0548 [0.0505, 0.0590]	<0.0001	20/20
0.327	-0.0012 [-0.0024, 0.0000]	0.0637	6/20	0.0682 [0.0637, 0.0727]	<0.0001	20/20

Table C3. Paired differences in conflict-detection AUC (exploratory; no equivalence margin was specified).

[56]Intervals are t-intervals over 20 seeds; see the note on p-values above.