Vagueness and Non-transitive Entailment

Paul Égré

Joint work with Robert van Rooij and Pablo Cobreros

Institut Jean-Nicod, CNRS

LoMOReVI - September 14, 2009

Tolerance and the sorites

Assume an indifference or similarity relation $\sim_P$:

[T] $\forall xy \in I : P(x) \wedge x \sim_P y \rightarrow P(y)$

Tolerance and the sorites

Assume an indifference or similarity relation $\sim_P$:

[T] $\forall xy \in I : P(x) \wedge x \sim_P y \rightarrow P(y)$

- Px: 'x is rich'; $x \sim_P y$: 'x's salary differs from y's by 1 euro'.

Tolerance and the sorites

Assume an indifference or similarity relation $\sim_P$:

[T] $\forall xy \in I : P(x) \wedge x \sim_P y \rightarrow P(y)$

- Px: 'x is rich'; $x \sim_P y$: 'x's salary differs from y's by 1 euro'.
- Interpret $x \sim_P y$ as $|x - y| \leq 1$ over the integers.

Tolerance and the sorites

Assume an indifference or similarity relation $\sim_P$:

[T] $\forall xy \in I : P(x) \wedge x \sim_P y \rightarrow P(y)$

- Px: 'x is rich'; $x \sim_P y$: 'x's salary differs from y's by 1 euro'.
- Interpret $x \sim_P y$ as $|x - y| \leq 1$ over the integers.
- Sorites: a contradiction follows from [T] and e.g. $\neg P(0)$ and $P(10^6)$.

Supervaluationism

- Partially defined predicates

Supervaluationism

- Partially defined predicates
- Truth relative to precisifications

Supervaluationism

- Partially defined predicates
- Truth relative to precisifications
- [T] is super-false: every precisification admits a cutoff, even if the cutoff can vary from one precisification to the other.

Supervaluationism

- Partially defined predicates
- Truth relative to precisifications
- [T] is super-false: every precisification admits a cutoff, even if the cutoff can vary from one precisification to the other.
- Sorites valid but not sound

Some issues

Problems:

- not so clear that supervaluationism gives a good account of the status of borderline cases.
- not clear that it can explain why the soritical premise is plausible

Alternative

- With fuzzy logicians: we agree that borderline cases are more likely to be ambivalent cases, cases at which P holds to some degree, and likewise for $\neg P$.

Alternative

- With fuzzy logicians: we agree that borderline cases are more likely to be ambivalent cases, cases at which P holds to some degree, and likewise for $\neg P$.
- Further agreement regarding the idea that a better account of the plausibility of tolerance is needed.

Alternative

- With fuzzy logicians: we agree that borderline cases are more likely to be ambivalent cases, cases at which P holds to some degree, and likewise for $\neg P$.
- Further agreement regarding the idea that a better account of the plausibility of tolerance is needed.
- Problem: prospects for a more qualitative approach, namely without anything as fine-grained as degrees of truth?

Subvaluationism

Hyde 1997

- Dual approach to supervaluationism

Subvaluationism

Hyde 1997

- Dual approach to supervaluationism
- Sub-truth= truth in at least one precisification

Subvaluationism

Hyde 1997

- Dual approach to supervaluationism
- Sub-truth= truth in at least one precisification
- *P* can be both true and false of borderline cases

Subvaluationism

Hyde 1997

- Dual approach to supervaluationism
- Sub-truth= truth in at least one precisification
- *P* can be both true and false of borderline cases
- Sorites invalid, but all premises can be sound, including some versions of tolerance.

Subvaluationism

Hyde 1997

- Dual approach to supervaluationism
- Sub-truth= truth in at least one precisification
- *P* can be both true and false of borderline cases
- Sorites invalid, but all premises can be sound, including some versions of tolerance.

Goal

- Explore a bivalent alternative to subvaluationism
- Instead of using partial models and precisifications: put more structure upon classical models
- Attempt to internalize the idea of tolerance: connect non-transitivity within structures to non-transitivity in entailment

Outline

- Subvaluationism

Outline

- Subvaluationism
- Tolerance and non-transitivity

Outline

- Subvaluationism
- Tolerance and non-transitivity

(Caveat: in progress)

Example

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#

Example

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1

Example

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

Paraconsistency

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

Paraconsistency

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

$M \models_{sbv} H(2)$
$M \models_{sbv} \neg H(2)$

Paraconsistency

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

$M \models_{sbv} H(2)$
$M \models_{sbv} \neg H(2)$
yet:

Paraconsistency

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

$M \models_{sbv} H(2)$
$M \models_{sbv} \neg H(2)$
yet:

$M \not\models_{sbv} H(2) \wedge \neg H(2)$

Validity and consequence

- $\models_{sbv} \phi$: ϕ is sub-true in every model M
- $\Gamma \models_{sbv} \psi$: every model that makes all ϕ in Γ sub-true makes ψ sub-true

Validity and consequence

- $\models_{sbv} \phi$: ϕ is sub-true in every model M
- $\Gamma \models_{sbv} \psi$: every model that makes all ϕ in Γ sub-true makes ψ sub-true

Proposition: $\models \phi \;\Leftrightarrow\; \models_{sbv} \phi$

Validity and consequence

- $\models_{sbv} \phi$: ϕ is sub-true in every model M
- $\Gamma \models_{sbv} \psi$: every model that makes all ϕ in Γ sub-true makes ψ sub-true

Proposition: $\models \phi \Leftrightarrow \models_{sbv} \phi$

However: $\Gamma \models \phi \not\Leftrightarrow \Gamma \models_{sbv} \phi$

Modus ponens

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

Modus ponens

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

- $H(1), H(1) \rightarrow H(2), H(2) \rightarrow H(3) \not\models_{sbv} H(3)$

Modus ponens

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

- $H(1), H(1) \rightarrow H(2), H(2) \rightarrow H(3) \not\models_{sbv} H(3)$
- modus ponens invalid

Modus ponens

	$H(3)$	$H(2)$	$H(1)$	$\neg H(1)$	$H(3) \rightarrow H(2)$	$H(2) \rightarrow H(1)$
M	1	#	0	1	#	#
M'	1	0	0	1	0	1
M''	1	1	0	1	1	0

- $H(1), H(1) \rightarrow H(2), H(2) \rightarrow H(3) \not\models_{sbv} H(3)$
- modus ponens invalid
- Sorites invalid in this form, and all conditional premises can be true together.

Observations

- $H(1) \rightarrow H(2), H(2) \rightarrow H(3) \not\models_{sbv} \forall n(H(n) \rightarrow H(n+1))$, even when $D_M = \{1, 2, 3\}$ for all M.

Observations

- $H(1) \rightarrow H(2), H(2) \rightarrow H(3) \not\models_{sbv} \forall n(H(n) \rightarrow H(n+1))$, even when $D_M = \{1, 2, 3\}$ for all M.
- $H(1) \wedge \forall n(H(n) \rightarrow H(n+1) \models_{sbv} H(3)$: in this form, the sorites reasoning remains valid.

Prospects

- Make tolerance valid in the form of a schema
- Internalize the idea of paraconsistency

Semi-orders

Luce 1956, van Rooij 2009

- xPy: "x is significantly greater/heavier/sweeter than y"

Semi-orders

Luce 1956, van Rooij 2009

- xPy: “x is significantly greater/heavier/sweeter than y”
- Structures (W, P) where P is a binary relation on W such that for all x, y, z, w:

Semi-orders

Luce 1956, van Rooij 2009

- xPy: "x is significantly greater/heavier/sweeter than y"
- Structures (W, P) where P is a binary relation on W such that for all x, y, z, w:

(I) $\neg xPx$

Semi-orders

Luce 1956, van Rooij 2009

- xPy: "x is significantly greater/heavier/sweeter than y"
- Structures (W, P) where P is a binary relation on W such that for all x, y, z, w:

(I) $\neg xPx$

(IO) if $xPy \wedge zPw$ then xPw or zPy

Semi-orders

Luce 1956, van Rooij 2009

- xPy: "x is significantly greater/heavier/sweeter than y"
- Structures (W, P) where P is a binary relation on W such that for all x, y, z, w:

(I) $\neg xPx$
(IO) if $xPy \wedge zPw$ then xPw or zPy
(ST) if xPy and yPz then xPw or wPz

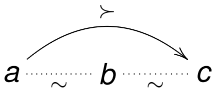
Semi-orders

Luce 1956, van Rooij 2009

- xPy: "x is significantly greater/heavier/sweeter than y"
- Structures (W, P) where P is a binary relation on W such that for all x, y, z, w:

(I)	$\neg xPx$
(IO)	if $xPy \wedge zPw$ then xPw or zPy
(ST)	if xPy and yPz then xPw or wPz

Indifference



- Let: $x \sim y$ iff neither xPy nor yPx.
- $\sim$ will be reflexive and symmetric, but typically non-transitive (Goodman 1951, Williamson 1994)

Interpretations based on semi-orders

Interpretations for first-order logic:

- Models are structures $M = \langle D, I, \sim \rangle$.

Interpretations based on semi-orders

Interpretations for first-order logic:

- Models are structures $M = \langle D, I, \sim \rangle$.
- $\sim$: function which, to each predicate P of the language, associates a relation $\sim_P$ over D that is reflexive and symmetric, possibly non-transitive.

Interpretations based on semi-orders

Interpretations for first-order logic:

- Models are structures $M = \langle D, I, \sim \rangle$.
- $\sim$: function which, to each predicate P of the language, associates a relation $\sim_P$ over D that is reflexive and symmetric, possibly non-transitive.
- Recursive definition of two satisfaction relations: classical ($\models$) and tolerant ($\models^t$)

Tolerance semantics (I)

Substitutional interpretation of quantification: every individual d has a name $\underline{d}$.

Tolerance semantics (I)

Substitutional interpretation of quantification: every individual d has a name $\underline{d}$.

$M \models^t P(\underline{a})$ iff $\exists d \sim_P a : M \models P(\underline{d})$

Tolerance semantics (I)

Substitutional interpretation of quantification: every individual d has a name $\underline{d}$.

$M \models^t P(\underline{a})$ iff $\exists d \sim_P a : M \models P(\underline{d})$

$M \models^t \neg\phi$ iff $M \not\models^t \phi$

$M \models^t \phi \wedge \psi$ iff $M \models^t \phi$ and $M \models^t \psi$

$M \models^t \forall x\phi$ iff for all $d \in D : M \models^t \phi[\underline{d}/x]$

Entailment relations

Let $[\![\phi]\!]^{(t)} := \{M : M \models^{(t)} \phi\}$.

Entailment relations

Let $[\![\phi]\!]^{(t)} := \{M : M \models^{(t)} \phi\}$.

- $\phi \models_t \psi$ iff $[\![\phi]\!] \subseteq [\![\psi]\!]^t$.

Non-transitivity

- $\models$ is transitive.

Non-transitivity

- $\models$ is transitive.

(1) If $P(\underline{a}) \models P(\underline{b})$ and $P(\underline{b} \models P(\underline{c})$, then $P(\underline{a}) \models P(\underline{c})$

Non-transitivity

- $\models$ is transitive.

(1) If $P(\underline{a}) \models P(\underline{b})$ and $P(\underline{b} \models P(\underline{c})$, then $P(\underline{a}) \models P(\underline{c})$

- $\models^t$ is not transitive

Non-transitivity

- $\models$ is transitive.

(1) If $P(\underline{a}) \models P(\underline{b})$ and $P(\underline{b} \models P(\underline{c})$, then $P(\underline{a}) \models P(\underline{c})$

- $\models^t$ is not transitive

(2) $P(\underline{a}) \models^t P(\underline{b})$ and $P(\underline{b}) \models^t P(\underline{c})$ are compatible with $P(\underline{a}) \not\models^t P(\underline{c})$

Non-transitivity

- $\models$ is transitive.

(1) If $P(\underline{a}) \models P(\underline{b})$ and $P(\underline{b} \models P(\underline{c})$, then $P(\underline{a}) \models P(\underline{c})$

- $\models^t$ is not transitive

(2) $P(\underline{a}) \models^t P(\underline{b})$ and $P(\underline{b}) \models^t P(\underline{c})$ are compatible with $P(\underline{a}) \not\models^t P(\underline{c})$

Consider models M such that $a \sim b$; $a \in I(P)$; $b, c \notin I(P)$; c is not $\sim$-related to any d in $I(P)$.

Is this good enough?

Let $M \models^{(t)} \underline{a} I_P \underline{b}$ iff $a \sim_P b$ in M, where every I_P is a distinguished predicate interpreted by $\sim_P$

Is this good enough?

Let $M \models^{(t)} \underline{a}I_P\underline{b}$ iff $a \sim_P b$ in M, where every I_P is a distinguished predicate interpreted by $\sim_P$

- Do we have: $\models^t \forall x \forall y (P(x) \wedge xI_Py \rightarrow P(y)$?

Is this good enough?

Let $M \models^{(t)} \underline{a} I_P \underline{b}$ iff $a \sim_P b$ in M, where every I_P is a distinguished predicate interpreted by $\sim_P$

- Do we have: $\models^t \forall x \forall y (P(x) \wedge x I_P y \rightarrow P(y))$?
- No: consider a model M with only three elements a, b, c such that $a \sim b \sim c$ but $a \nsim b$; $I(P) = \{a\}$. $M \models^t P(\underline{b})$, however $M \not\models^t P(\underline{c})$.

Is this good enough?

Let $M \models^{(t)} \underline{a} I_P \underline{b}$ iff $a \sim_P b$ in M, where every I_P is a distinguished predicate interpreted by $\sim_P$

- Do we have: $\models^t \forall x \forall y (P(x) \wedge x I_P y \rightarrow P(y))$?
- No: consider a model M with only three elements a, b, c such that $a \sim b \sim c$ but $a \nsim b$; $I(P) = \{a\}$. $M \models^t P(\underline{b})$, however $M \not\models^t P(\underline{c})$.

Consequence: the tolerance principle is not t-valid.

Tolerant semantics (II)

- $P(a) \models^t P(a)$ but obviously $\neg P(a) \not\models^t \neg P(a)$

Tolerant semantics (II)

- $P(a) \models^t P(a)$ but obviously $\neg P(a) \not\models^t \neg P(a)$
- Hence: weaken the meaning of negation
- Idea: define two relations $\models^t$ and $\models^s$ in terms of each other.

t-satisfaction

$M \models^t P(\underline{a})$ iff $\exists d \sim_P a : M \models P(\underline{d})$

$M \models^t \neg\phi$ iff $M = |^s \phi$

$M \models^t \phi \wedge \psi$ iff $M \models^t \phi$ and $M \models^t \psi$

$M \models^t \forall x \phi$ iff for all $d \in D : M \models^t \phi[\underline{d}/x]$

t-satisfaction

$$M \models^t P(\underline{a}) \text{ iff } \exists d \sim_P a : M \models P(\underline{d})$$

$$M \models^t \neg\phi \text{ iff } M =|^s \phi$$

$$M \models^t \phi \wedge \psi \text{ iff } M \models^t \phi \text{ and } M \models^t \psi$$

$$M \models^t \forall x\phi \text{ iff for all } d \in D : M \models^t \phi[\underline{d}/x]$$

$$M =|^t P(\underline{a}) \text{ iff } \forall d \sim_P a : M \not\models P(\underline{d})$$

$$M =|^t \neg\phi \text{ iff } M \models^s \phi$$

$$M =|^t \phi \wedge \psi \text{ iff } M =|^t \phi \text{ or } M =|^t \psi$$

$$M =|^t \forall x\phi \text{ iff } \exists d \in D : M =|^t \phi[\underline{d}/x]$$

s-satisfaction

$M \models^s P(\underline{a})$ iff $\forall d \sim_P a : M \models P(\underline{d})$

$M \models^s \neg\phi$ iff $M = |^t \phi$

$M \models^s \phi \wedge \psi$ iff $M \models^s \phi$ and $M \models^s \psi$

$M \models^s \forall x \phi$ iff for all $d \in D : M \models^s \phi[\underline{d}/x]$

s-satisfaction

$M \models^s P(\underline{a})$ iff $\forall d \sim_P a : M \models P(\underline{d})$

$M \models^s \neg\phi$ iff $M =|^t \phi$

$M \models^s \phi \wedge \psi$ iff $M \models^s \phi$ and $M \models^s \psi$

$M \models^s \forall x\phi$ iff for all $d \in D : M \models^s \phi[\underline{d}/x]$

$M =|^s P(\underline{a})$ iff $\exists d \sim_P a : M \not\models P(\underline{d})$

$M =|^s \neg\phi$ iff $M \models^t \phi$

$M =|^s \phi \wedge \psi$ iff $M =|^s \phi$ or $M =|^s \psi$

$M =|^s \forall x\phi$ iff $\forall d \in D : M =|^s \phi[\underline{d}/x]$

Consequence

- $\models^t \phi$ iff for all M, $M \models^t \phi$
- $\phi \models^t \psi$ iff $[\![\phi]\!] \subseteq [\![\psi]\!]^t$
- $\phi \models^s \psi$ iff $[\![\psi]\!]^s \subseteq [\![\phi]\!]$

Tolerance

$$\models^t \forall x \forall y (P(x) \wedge x I_P y \rightarrow P(y))$$

Proof:
$\Leftrightarrow$ for all a, b in M:
$M =|^s P(\underline{a})$ or $M =|^s \underline{a} I_P \underline{b}$ or $M \models^t P(\underline{b})$.

Suppose: $a \sim_P b$.
Either $a \in I(P)$: then $M \models^t P(\underline{b})$
Or: $a \notin I(P)$: then trivially $M =|^s P(\underline{a})$.

The sorites

$$1 \overset{}{\underset{\sim}{\cdots\cdots}} 2 \underset{\sim}{\cdots\cdots} 3 \underset{\sim}{\cdots\cdots} 4$$

$$H \qquad H \qquad \neg H \qquad \neg H$$

$M \models^t H(1)$
$M \not\models^t H(4)$
$M \models^t H(1) \rightarrow H(2)$
$M \models^t H(2) \rightarrow H(3)$
$M \models^t H(3) \rightarrow H(4)$

Paraconsistency

$P(\underline{a}) \wedge \neg P(\underline{a})$ is t-satisfiable whenever a is similar to some d in P, and to some d' not in P, namely whenever a is a borderline point.

- Yet: $\models^t P(a) \vee \neg P(a)$

Departure from classical logic

$$\models (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$$

Departure from classical logic

$\models (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$

However:

Departure from classical logic

$$\models (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$$

However:

$$\not\models^{t} (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$$

Departure from classical logic

$\models (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$

However:

$\not\models^t (P(\underline{a}) \wedge P(\underline{a}) \rightarrow P(\underline{b})) \rightarrow P(\underline{b})$

Proof: suppose a is a borderline case, such that $M \models^t P(\underline{a}) \wedge \neg P(\underline{a})$.
Then $M \models^t P(\underline{a}) \wedge (\neg P(\underline{a}) \vee P(\underline{b}))$
But it can be that $M \not\models^t P(\underline{b})$ (if b is not similar to any P-element).

Summary and comparisons

The second version of $\models^t$:

- non-transitive consequence relation ($\neq$ subvaluationism)
- validates the inductive premise of the sorites ($\neq$ subvaluationism)
- paraconsistent: some classical contradictions are satisfiable ($\neq$ subvaluationism)
- loses some classical validities ($\neq$ subvaluationism)

Further work needed

- Characterize t-validities
- Axiomatize t-entailment
- See how far we then get from CL
- Further connections to other frameworks?