Lecture 4: Some alternative approaches

Marco Degano Robert van Rooij

Formal Theories of Vagueness
ESSLLI 2026, Prague

6 August 2026

Readings

- Williamson, T. (2002). *Vagueness*. Routledge.
- Cobreros, P., Egré, P., Ripley, D., van Rooij, R. (2015). *Vagueness, Truth and Permissive Consequence*. In: Achourioti, T., Galinon, H., Martínez Fernández, J., Fujimoto, K. (eds) Unifying the Philosophy of Truth. Logic, Epistemology, and the Unity of Science, vol 36. Springer, Dordrecht.

Outline

1. Epistemicism

2. Contextualism

3. Pragmatic approach

4. ST: Non-transitive logic

5. Empirical support and Pragmatics

The epistemic solution

Timothy Williamson

- **Sharp but unknowable.** Vague expressions have *sharp* extension/cutoffs. Our ignorance makes them appear vague.
- There is a precise number distinguishing *bald* from *not bald*, but **we do not know** it.
- Vagueness = **epistemic limitation**, not semantic indeterminacy.

Inexact knowledge & margins of error

- (Williamson, *Vagueness* 1994) Knowledge is **safe**: it must be stable under small changes.
- **Margin-for-error schema** (for a vague predicate P and metric d):
$$Know_\alpha(P(x)) \;\Rightarrow\; \forall y\big(d(x,y) \leq \alpha \;\Rightarrow\; P(y)\big)$$
- Intuition: if you *know* that x is P, then anything α-close to x must also be P. Near the sharp cutoff, this *forces* ignorance: neither *know P* nor *know not P* can hold.

Fixed margin models

A fixed-margin (Kripke) model $M = \langle W, R, V \rangle$ with metric d and error parameter $\alpha > 0$:

$$R(x, y) \text{ iff } d(x, y) \leq \alpha$$

Reading: $R(x, y)$ = "y is within the α-margin of x"

Assume for all $x, y, z \in W$:

$$d(x, y) = 0 \Leftrightarrow x = y, \qquad d(x, y) = d(y, x), \qquad d(x, z) \leq d(x, y) + d(y, z).$$

Consequences for R:

- **Reflexive** ($d(x, x) = 0 \leq \alpha$)
- **Symmetric**
- **Not necessarily transitive** ($d(x, z)$ may be 2α)

[Égré and Bonnay 2010; Williamson 2002]

Knowledge operator

$$M, x \models \Box\varphi \quad \text{iff} \quad \forall y\ (R(x, y) \Rightarrow M, y \models \varphi)$$

- $\Box\varphi$ at x: "I **know** φ throughout the α-neighborhood of x".
- We work with a simple line model: $W = \mathbb{N}$, $d(m, n) = |m - n|$, $\alpha = 1$.

[Égré and Bonnay 2010]

Example (line model, $\alpha = 1$)

$$\begin{tikzcd}
0 \arrow[loop above] \arrow[r, leftrightarrow] & 1 \arrow[loop above] \arrow[r, leftrightarrow] & 2 \arrow[loop above] \arrow[r, leftrightarrow] & 3 \arrow[loop above] \arrow[r, leftrightarrow] & 4 \arrow[loop above] \arrow[r, leftrightarrow] & 5 \arrow[loop above] \arrow[r] & \cdots \\
p & p & p & p & \neg p & \neg p &
\end{tikzcd}$$

Take p as a vague predicate (e.g., “not a heap”). So small n satisfy p, large n satisfy $\neg p$.

- At 2: $\Box p$ holds (both $1, 2, 3$ satisfy p). At 5: $\Box\neg p$ holds.
- At the boundary $3, 4$: $\neg\Box p \wedge \neg\Box\neg p$ (ignorance).

Unknown sharp cutoff

- There is a **sharp cutoff** n^*. For $n < n^*$: p (not a heap). For $n \geq n^*$: $\neg p$ (heap).
- **Margin-for-error** forbids knowledge at the α-neighbors of n^*.
- At a clear case like 0, we naturally want very strong epistemic security about p: ideally $\Box p, \Box^2 p, \Box^3 p, \ldots$ all hold.
- Fixed-margin models block this ideal: they validate $\Box p$ at clear cases but *do not* generally validate all iterations $\Box^n p$. Epistemicist reply: as we iterate the knowledge operator, knowledge **"erodes"**.

Knowledge axioms

(1) $\Box\varphi \to \varphi$

Factivity: if I know φ, then φ is true.

(2) $\Box\varphi \to \Box\Box\varphi$

Positive introspection: if I know φ, then I know that I know φ.

(3) $\neg\Box\varphi \to \Box\neg\Box\varphi$

Negative introspection: if I don't know φ, then I know that I don't know φ.

- In the fixed-margin semantics, (1) is valid: knowledge is factive.
- If we also require (2) or (3) to be valid at all worlds, then the accessibility relation R must be **transitive**.
- But non-transitivity of R is exactly what allows the model to block the Sorites.
  So, on the epistemicist picture, (2) and (3) are *not* generally valid for inexact knowledge.

Tolerance and 2-dimensional semantics

- **Tolerance**: $\forall x, y((\Box Px \wedge R_P(x, y)) \rightarrow Py)$
  standard tolerance assumption is **weakened**
  because stronger antecedent

- **Accessible** epistemic **worlds** w.r.t P:
  about the cutt-off point of P

Worlds are accessible from w only if the **meaning** of P does not change a lot compared to w.

Like in 2-dimensional semantics, worlds play double role: gives meaning (cut-off point for 'Tall') + facts (John's height)

Assessing the epistemic response

- **Counterintuitive sharpness:** It posits **precise** cutoffs for *tall*, *heap*, etc.
- **Ignorance challenge:** If there *is* a cutoff, why can't competent speakers know it?
- **Higher-order ignorance:** Do we *know* the margin α? If not, do margins themselves admit margins? (Iterated ignorance.)

Outline

1. Epistemicism

2. Contextualism

3. Pragmatic approach

4. ST: Non-transitive logic

5. Empirical support and Pragmatics

Contextualism

- Vague predicates are **context-sensitive** (comparison class / standard shifts).
  Whether John is tall depends one who he is compared to.

- The conditional premises in Sorites are *intuitively compelling*.
  Contextualism explains this pull via **context shifts**.

Indistinguishability Relation I

We can state the Sorites premise as:

If x is P and x is indistinguishable from y, then y is P.

$$(P(x) \;\wedge\; xI_Py) \;\rightarrow\; P(y)$$

What properties should I_P have? In particular, can I_P be *transitive*?

From "significant difference" $\succ_P$ to indistinguishable I_P

Define "significantly P-er than" by $x \succ_P y$. Then set:

$$xI_Py \quad := \quad \neg(x \succ_P y) \ \wedge \ \neg(y \succ_P x)$$

- If $\succ_P$ is a *strict weak order* (irreflexive, transitive, almost-connected)[1], then I_P is an **equivalence relation** (check this as an exercise). Hence transitive.
- To avoid transitivity, we use a more realistic **"just noticeable difference"** ordering.

[1] **Almost connectedness**: $\forall x \, \forall y \, \forall z \, \big(R(x,y) \rightarrow (R(x,z) \ \vee \ R(z,y))\big)$.

Semi orders

We define $\succ_P$ as a **semi-order**:

Irreflexive: $\forall x : \neg(x \succ_P x)$
Interval-order: $\forall x, y, v, w : (x \succ_P y \wedge v \succ_P w) \rightarrow (x \succ_P w \vee v \succ_P y)$
Semi-transitive $\forall x, y, z, v : (x \succ_P y \wedge y \succ_P z) \rightarrow (x \succ_P v \vee v \succ_P z)$

- Irreflexive: nothing is *noticeably* more P than itself.
- Interval-order: if we can tell x is better than y and v is better than w, then at least one "better" element also beats the other pair's "worse" element (there are no two completely independent just-noticeable differences)
- Semi-transitive: if x is better than y and y is better than z, then any fourth option v must line up with one of the extremes, x or z

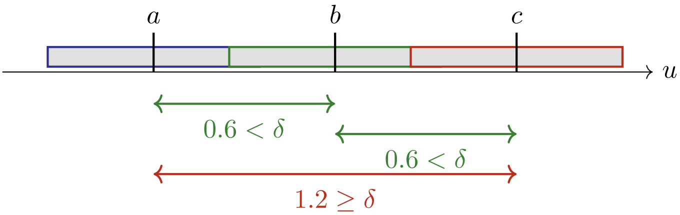
Hence I_P is reflexive and symmetric, but need not be transitive.

Semi orders: (non-transitive I_P)

Choose $\delta = 0.7$ and $u(a) = 0,\ u(b) = 0.6,\ u(c) = 1.2$.

$$|u(a) - u(b)| = 0.6 < \delta, \quad |u(b) - u(c)| = 0.6 < \delta, \quad |u(a) - u(c)| = 1.2 \geq \delta.$$

Thus $a\, I_P\, b$ ✓, $b\, I_P\, c$ ✓, but $a\, I_P\, c$ ✗.



Width-δ windows overlap for neighboring pairs, but overlap does not chain from a to c.

Context-dependent I

Contextualist move: I_P is **context-dependent** and the context *changes* along a Sorites sequence.

1. Start: Px_1 and consider $x_2 : x_1 I_P x_2$. By Tolerance, also Px_2
2. Consider x_3. Add x_3 to context c, and check whether $\boxed{x_2 I_P^c x_3}$
3. If so, conclude by Tolerance that also Px_3, etc.

Similarity relativized to a comparison class $c \subseteq D$:

$$xI_P^c y \quad \text{iff} \quad \forall z \in c :\ xI_P z \leftrightarrow yI_P z$$

"x and y are not (even indirectly) distinguishable relative to c."

but can be that $xI_P y$ and $yI_P z$, but $x \succ_P z$. So, not $yI_P^{\{x,y,z\}} z$

General: I_P^c is an **equivalence relation**, and thus also transitive!

Local validity vs. global invalidity

Tolerance principle is safe *in isolation* at its own context c:

$$\big(P(x,c)\ \wedge\ xI_P^c y\big)\ \rightarrow\ P(y,c)$$

where here $c = \{x, y\}$.

But we cannot conjoin premises across *different* contexts:

(1) $P(x,c)$ with $c = \{x, y, z\}$
(2) $\big(P(x,c) \wedge xI_P^c y\big) \rightarrow P(y,c)$ with $c = \{x, y\}$
(3) $\big(P(y,c) \wedge yI_P^c z\big) \rightarrow P(z,c)$ with $c = \{y, z\}$
(4) $P(z,c)$ with $c = \{x, y, z\}$

From (1)-(3) *you cannot derive* (4): we would need (2) and (3) on $c = \{x, y, z\}$

Contextualism

Experimental angle:

- Forced-march tasks (stepwise “still P?”) often show tolerance
- Presentation/manipulation of *comparison class* and *stimulus variation* shift judgments.

Philosophical concerns:

- Equivocation worry: Are we merely “changing the subject”?
- Fixing c: Even if we stipulate a fixed c, the paradox can still feel compelling: what explains that pull?
- Higher-order vagueness: Standards themselves seem vague. Can contextualism iterate the story?

Outline

1. Epistemicism
2. Contextualism
3. Pragmatic approach
4. ST: Non-transitive logic
5. Empirical support and Pragmatics

Pragmatic approach

Idea: The use of vague predicates (and Tolerance reasoning) is only allowed in **appropriate contexts**

Appropriate context: where the Sorites reasoning does not arise
$\Rightarrow$ a context where there is a **gap** between P and $\neg P$ individuals

$\forall x, y((P(x,c) \wedge xI_P y) \rightarrow P(y,c))$, for all appropriate c

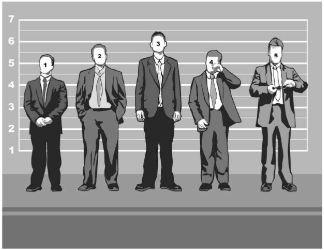
$Appropriate(c)$ iff
$\neg\exists x_1 \cdots x_n \in c : P(x,c) \wedge \neg P(x_n,c) \wedge x_1 I_P x_2 \wedge \cdots \wedge x_{n-1} I_P x_n$

Rayo, van Rooij ($\approx$ 2010), Pagin (most sophisticated), arguably all based on Wittgenstein II

Problems pragmatic approach

- How to account for context change in the middle of a conversation? From appropriate to inappropriate. Would we retract earlier claims?
- Even in an **in**appropriate context, it seems that we judge the tallest man to be tall (and the shortest to be short)
- And we have judgements about middlemen as well ....

Vagueness and Experiments



- Alxatib & Pelletier (2011): 5 men of increasing height.
- Participants judged, for each man:
  - "He is tall"
  - "He is not tall"
  - "He is tall and not tall"
  - "He is neither tall nor not tall"
- Results for man #2:
  - "He is **tall and not tall**": True 44.7%, False 40.8%, Can't tell 14.5%
  - "He is **neither tall nor not tall**": True 53.9%, False 42.1%, Can't tell 4.0%.

Outline

1. Epistemicism

2. Contextualism

3. Pragmatic approach

4. ST: Non-transitive logic

5. Empirical support and Pragmatics

Tolerance and Vagueness

- Measurement Tolerance: $\forall x, y : (x \sim_P y \rightarrow (x \not\succ_P y \wedge y \not\succ_P x)$

- Kamp (1981), Wright, Dummett:

  What makes a predicate vague, is that it is tolerant

- Tolerance principle: **[P]** $\forall x, y : (Px \wedge x \sim_P y) \rightarrow Py$

- Tolerance is **constitutive** of its meaning.

Solutions to the Sorites

All: give up on $[\mathbf{P}]$ $\quad \forall x, y \in I : (P(x) \wedge x \sim y) \rightarrow P(y)$

- Supervaluation: $[\mathbf{P}]$ false,
  but no particular instantiation (super)true.
- Other solutions: *weakening* of $[\mathbf{P}]$ is true.
  1. Fuzzy logic: $[\mathbf{P}]$ *almost* true
  2. Three-valued logic
  3. Epistemics
  4. Contextual solution
  5. Pragmatic solution
- None of them takes tolerance seriously (enough)!

Tolerant interpretation

Start with FOL model, with $\sim_P$ a non-transitive similarity relation

- $M \models^c \phi$ defined in the usual way.
- $\boxed{M \models^t P(\underline{a})}$ iff $\boxed{\exists d \sim_P a}$: $\boxed{M \models^c P(\underline{d})}$ with $\underline{d}$ name for 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 $d \in I_M$: $M \models^t \phi[^x/_{\underline{d}}]$.

- $\underline{[[P a]]}^c \subset [[P\underline{a}]]^t$ $\rightsquigarrow$ Tolerant truth **weaker**

- $[[\neg P\underline{a}]]^c \supset [[\neg P\underline{a}]]^t$ $\rightsquigarrow$ Tolerant truth **stronger**

Satisfaction: TCS (JoPL)

- $M \models^t P(\underline{a})$ iff $\exists d \sim_P a$: $M \models^c P(\underline{d})$
- $M \models^t \neg\phi$ iff $M \not\models^s \phi$
- $M \models^t \phi \wedge \psi$ iff $M \models^t \phi$ and $M \models^t \psi$
- $M \models^t \forall x \phi$ iff for all $d \in D$: $M \models^t \phi[\underline{d}/x]$

- $M \models^s P(\underline{a})$ iff $\forall d \sim_P a$: $M \models^c P(\underline{d})$
- $M \models^s \neg\phi$ iff $M \not\models^t \phi$
- $M \models^s \phi \wedge \psi$ iff $M \models^s \phi$ and $M \models^s \psi$
- $M \models^s \forall x \phi$ iff for all $d \in D$: $M \models^s \phi[\underline{d}/x]$

- $\Gamma \models^{st} \psi$ iff $\forall M, \forall \phi \in \Gamma : M \models^s \phi \Rightarrow M \models^t \psi$

Cobreros, Egre, Ripley, van Rooij (2012), ‘Tolerant, Classical, Strict, JPL

Properties

- $[[\phi]]^s \subseteq [[\phi]] \subseteq [[\phi]]^t$ $\quad\rightsquigarrow\quad$ Also for $\neg Pa$
- $Pa \wedge \neg Pa$ not a tolerant contradiction
- $Pa \vee \neg Pa$ not a strict tautology
- Connection with LP and K3?

Three-valued

A ST-model is a three-valued model M, of the sort familiar from Strong Kleene logic or LP. The domain must include $\mathcal{L}$.

Where I_M is an interpretation:

- $I_M(A) \in \{1, 1/2, 0\}$
- $I_M(\top) = 1; I(\bot) = 0$
- $I_M(A \wedge B) = \min(I_M(A), I_M(B))$
- $I_M(\neg A) = 1 - I_M(A)$
- $I_M(\forall x A) = \min(\{I'_M(A) : I'$ is an x-variant of $I_M\})$

- $\Gamma \models^{st} \psi$ iff $\forall M, \forall \phi \in \Gamma : I_M(\phi) = 1 \Rightarrow I_M(\psi) \neq 0$

Strict vs. tolerant truth

Fix the Strong Kleene truth-functions on $\{0, i, 1\}$.

- **Strict truth** (s): $v(\varphi) = 1$.
- **Tolerant truth** (t): $v(\varphi) \neq 0$ (i.e. 1 or i).
- Duality: φ is t-true iff $\neg\varphi$ is not s-true (and vice versa).
- “strict” tracks full truth
- “tolerant” tracks non-falsity (or permissive assertability).

Mixed consequence (four variants)

For $n, m \in \{\mathsf{s}, \mathsf{t}\}$:
$\Gamma \models_{nm} \varphi$ iff there is no Strong Kleene valuation v with $[\forall \gamma \in \Gamma,\ v \models_n \gamma]\ \&\ v \not\models_m \varphi$.

- $\models_{\mathsf{ss}}$ = K3 (preserves 1).
- $\models_{\mathsf{tt}}$ = LP (preserves $\neq 0$).
- $\models_{\mathsf{ts}}$ is *empty*[2] (premises weaker than conclusions).
- $\models_{\mathsf{st}}$ = **ST**: from strict premises to tolerant conclusions.

$$\Gamma \models_{ST} \varphi \text{ iff } [\forall \gamma \in \Gamma,\ v(\gamma) = 1] \Rightarrow v(\varphi) \neq 0$$

[2] for a constant-free language

Conditionals and why ST keeps Modus Ponens

$$v(A \to B) = \max(1 - v(A),\ v(B))$$

$$v(A \to B) \neq 0 \quad \Leftrightarrow \quad [v(A) \neq 1] \text{ or } [v(B) \neq 0].$$

So the object-language $\to$ mirrors the meta-pattern s⇒t:

if A is strictly true, then B is tolerantly true.

- ST validates **Modus Ponens** and the **Deduction Theorem.**

Classicality of ST (base language)

Theorem (ST = Classical Consequence)

For any Γ, φ in the base language,

$$\Gamma \models_{\mathrm{CL}} \varphi \text{ iff } \Gamma \models_{\mathrm{ST}} \varphi$$

Prove the contrapositive:

- ($\Leftarrow$) Any classical countermodel is an ST countermodel.
- ($\Rightarrow$) Any ST countermodel can be *refined* by replacing each $\frac{1}{2}$ with 0 or 1 so as to yield a classical countermodel.

(Write out the proof in full: you may use a refinement argument from SK to classical valuations, which you need to prove by induction on formulas.)

Adding vagueness: ST with indifference (STVP)

- We extend the language with a crisp similarity predicate I_P for each vague P. With the *closeness proviso* on all valuations v:

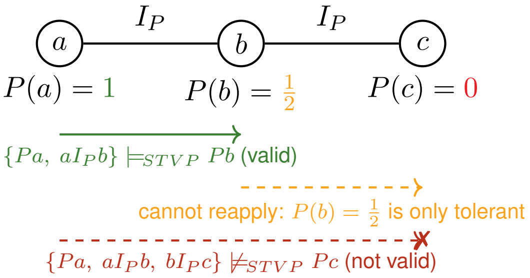
$$v \models_{\mathsf{s}} aI_Pb \quad \text{iff} \quad v \models_{\mathsf{t}} aI_Pb \quad \text{iff} \quad |v(Pa) - v(Pb)| < 1$$

- Hence I_P is reflexive symmetric, but *not necessarily transitive*.
- **Tolerance (valid in STVP):**

$$Pa,\, aI_Pb \models_{STVP} Pb \qquad\qquad \models_{STVP} \forall xy\,(Px \wedge xI_Py \rightarrow Py)$$

Intuition: if Pa holds strictly and b is P-indistinguishable from a, then Pb holds at least tolerantly.

STVP: non-transitive consequence (Sorites blocked)



- STVP validates *local* Tolerance, but the tolerant conclusion Pb cannot serve as the strict premise of the next step. Hence consequence is *non-transitive* and the Sorites chain fails.

ST and the Sorites

- Keeps the intuitive *Tolerance* principle in conditional form.
- Blocks the paradox via *non-transitive* consequence once I_P is present.
- Preserves classical behavior (Modus Ponens, Deduction Theorem, ...) on the base language.
- but tolerance principle cannot be strictly true

Note: one can use **ST** logic also based on other semantic bases:

1. Weak Kleene: **ST** logic is still classical logic
2. Super and Sub-valuationism (Cobreros et al. 2013)
3. Fuzzy logic: $\Gamma \models \phi$ iff $\forall v, \forall \gamma \in \Gamma : v(\gamma) > 0.5$, then $v(\phi) \geq 0.5$
4. Intuitionistic logic (Fitting)
5. etc.

ST and Evans' problem of Vague Identity

Assume $I(\nabla(\underline{a} = \underline{b})) = 1$ iff $I(\underline{a} = \underline{b}) = \frac{1}{2}$

1. $\nabla(\underline{a} = \underline{b})$
2. $\lambda x(\nabla \underline{a} = x)\underline{b}$ Lambda abstraction
3. $\neg\lambda x(\nabla \underline{a} = x)\underline{a}$
4. $\underline{a} \neq \underline{b}$ Leibniz' Law
5. $\neg\nabla(\underline{a} = \underline{b})$

- all steps are ST-valid
- $2, 3 \models^{st} 4 \quad 4 \models^{st} 5,$ because $I(\underline{a} \neq \underline{b}) = \frac{1}{2}$
- but $1 \not\models^{st} 5$

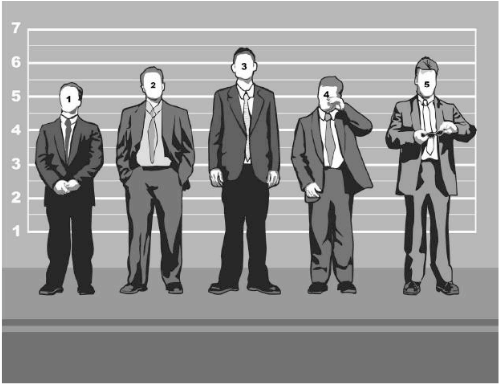
Outline

1. Epistemicism
2. Contextualism
3. Pragmatic approach
4. ST: Non-transitive logic
5. Empirical support and Pragmatics

Pragmatic analysis of Penumbral Connections

- Fine: Global quantification crucial for Penumbral connections
- Indeed, our analysis non-standard for $Pa \wedge \neg Pa \quad Pa \vee \neg Pa$
- But we don't see why this is problematic (see below)
- But if a slightly shorter than $b \Rightarrow M \models^t Ta \wedge \neg Tb$ is possible!!
- Pragmatics: always interpret as strongly as possible $\Rightarrow$ strictly
- Strictly interpretation of $Pa \wedge \neg Pb \Rightarrow a >_P b$
  But this is impossible if a slightly shorter than b

Use of contradictions (Alxatib & Pelletier, Ripley)



- V: Is 2 tall? A: no. V: Is 2 not tall? A: no.
- V: is 2 tall and not tall? A: yes.

Experiment 1: Ripley

Contradictions for borderline cases:

1. The circle is near the square and isn't near the square
2. The circle both is and isn't near the square
3. The circle neither is near the square nor isn't near the square
4. The circle neither is nor isn't near the square

All agreed with to same amount

Contradictions and excluded middle

Data from Ripley (2009), Alxatib and Pelletier (2009), Serchuk, Hargreaves and Zach (2008) indicate that:

- Subjects are more willing to accept $P(x) \wedge \neg P(x)$ for borderline cases (Ripley 2009, Alxatib and Pelletier 2009)
- Subjects are less willing to accept $P(x) \vee \neg P(x)$ for borderline cases (Serchuk et al. 2008)

The data agree with the strongest meaning hypothesis.

Problem: 'John is tall and not tall, **or** Mary is rich'

Solution (Cobreros et al, 2015) for this made use of truth-makers.