Vague Classes and a Resolution of the Bald Paradox

BEIHAI ZHOU AND LIYING ZHANG

ABSTRACT. In this paper, we introduce the concept of a vague class, based on which we give a resolution for the bald paradox, and explain how it arises. In a formalization part, we add equality symbols $\approxeq_P$ and monadic predicates Θ_P to a first-order language. Based on $\approxeq_P$ and Θ_P, vague predicates are formalized in an extended first-order language $\mathcal{L}^*$, and then vague classes are captured in a first-order way. Capturing vague predicates (vague classes) in an extended first-order language is a new method for dealing with vagueness, different from approaches that try to capture vagueness by introducing more truth values.[1]

1 Motivation

Vagueness is a common phenomenon in natural language. The literature on vagueness flourished since the 1970s, and there are different proposals to interpret vagueness and resolve the Sorites paradox. Among these approaches, three-valued logics, fuzzy theories, super-valuationism and epistemicism are the most influential. Though influential, each approach has been questioned from a technical aspect or from its intuitive reasonability.

This paper focuses on the simplest and typical vague expressions such as bald, tries to give a concept of vague class, and based on vague classes, a resolution to the Bald Paradox is given. An adequate study of vagueness needs to answer three questions: (1) What are the nature of vague expressions? (2) What about the truth value of borderline cases? (3) Why is the conclusion of the Bald Paradox wrong, and why does the reasoning look so reasonable? This paper tries to answer these questions. Section 2 is about Questions (1) and (2), Section 3 is about (3), and Section 4 continues with the analysis in Section 2.

In this paper, "bald" and the Bald Paradox are the running example. In fact, "bald" only represents one type of vague class, but a simplest one. Vague classes as defined in this paper, based on the case of "bald", cannot cover all other types of vagueness. In order to cover more types, we need further study.

[1]The paradox to be discussed is known under various names in the literature, such as the Sorites, or the Paradox of the Bald Man. In this paper, for brevity, we use the term "Bald Paradox", emphasizing our focus on the functioning of the vague predicate bald.

2 Vague classes

2.1 The intuition of a vague class

In fuzzy logic, there are fuzzy sets corresponding to vague expressions. The core concept for fuzzy sets is *degree of membership*. A member does not simply belong to a set or not, there is a measures to how an object belongs to a set. For instance, for "young people", we may use age as a criterion for membership; for "bald", we may use the number of hairs. Membership can be established by 'scientific' methods.

Degree of membership depends on a measure perspective. But in our view, understanding a vague predicate in daily life usually involves a natural perspective of agents' impressions, the scientific measure perspective is something acquired.[2] So, we will introduce the concept of vague class based on impressions. Classes with unclear bounds are *vague classes*, for instance, *bald*, *young*, *fatty*, *good student*, etc.

Consider the formation of the vague class for "bald". First, there are samples of objects that we consider bald. Second, if an individual is indistinguishable from a sample with respect to being bald, then that individual is bald. Hence, a vague class has two basic elements: the *sample* and the *distinguishing power*.

For distinguishing power, we use an indistinguishability relation. Through indistinguishability surrounding the samples, some objects are gathered, forming an object class. The higher the distinguishing power, the fewer indistinguishable objects left. Classes obtained in this way come wholly from an agent's cognition. Because our distinguishing power is not stable, this kind of class is vague, it is influenced by many factors, such as context and situation, and different cognitive agents may have different distinguishing power.

Something to be highlighted is that whether an individual is bald or not depends on our overall impression. In the process of forming this impression, we did not count numbers of hairs. In other words, "bald" is not a class obtained by counting numbers of hair, it is obtained by our cognitive impression. Although the number of hairs is relevant to determining whether someone is bald or not (more than a certain amount is surely non-bald, less than a certain amount is surely bald), since adding or removing several hairs does not affect the overall impression, the number of hairs is not an exact factor for determining whether someone is bald or not.

Here, two kinds of cognitive styles are involved. One is to get an overall visual impression with 'distance', the other is a closer observation of the number of hairs. For simplicity, we call these two ways as 'macro-impression' and 'micro-count' respectively. "Bald" is a class obtained by macro-impression, not by micro-count.

[2] For the details of 2 perspective, see 3.1.

2.2 Mathematical description of vague classes

A vague class has two basic elements: *sample* and *indistinguishability relation*.

Sample Samples are standard individuals confirmed by an agent. Samples are always about something, This something is the sample's subject. For example, there are samples of "bald". We denote the subject of a sample as τ, and the sample set of subject τ as θ_τ.

When we say that two objects are indistinguishable, the concept of an indistinguishability relation is needed. We introduce two kinds of indistinguishability relation in this section.

Indistinguishability relation 1: $x \approx_\tau y$ (x and y are indistinguishable on τ).

This says that, in a macroscopic perspective, individuals are indistinguishable on some subject. For example, in a macroscopic perspective, x and y are indistinguishable on the subject "bald", or on the subject "tall". It is possible that x and y are indistinguishable on the subject "bald", but distinguishable on the subject "tall". Being indistinguishable on some subject does not mean being indistinguishable on all subjects.

x and y being indistinguishable on τ can also be understood as: on τ, the differences of x and y are invisible (cannot be detected), or the difference is so tiny that it can be ignored.

Indistinguishability relation 2: $x \approx_{(\tau,\varphi)} y$

In indistinguishability relation 2, a distinguishable condition or circumstance φ is added: under condition φ, on subject τ, x and y are indistinguishable.

This relation will be denoted as: $\varphi \rightarrow x \approx_\tau y$ or $x \approx_{(\tau,\varphi)} y$.

Since whether two individuals are distinguishable or not depends on an agent's cognitive ability, there are many influencing factors. For example, to determine whether a man is bald or not is influenced by viewing distance, lighting, sum of hairs, colors, etc. All these factors are denoted by ϕ.

Definition 1 An indistinguishable set on τ under condition φ and based on a sample x is: $\phi_{(\tau,\varphi)}(x) = \{y : x \approx_{(\tau,\varphi)} y\}$.

Definition 2 Let θ_τ be a sample set with subject τ, φ is a distinguish condition.

Vague class is: $A(\theta_\tau, \varphi) = \bigcup_{x\in\theta_\tau} \phi_{(\tau,\varphi)}(x)$.

The *core* of $A(\theta_\tau, \varphi)$ is: $\kappa(A(\theta_\tau, \varphi)) = \bigcap_{x\in\theta_\tau} \phi_{(\tau,\varphi)}(x)$

The *border area* of $A(\theta_\tau, \varphi)$ is:

$\mu(A(\theta_\tau, \pi, \varphi)) = A(\theta_\tau, \varphi) - \kappa(A(\theta_\tau, \varphi)) = \bigcup_{x\in\theta_\tau} \phi_{(\tau,\varphi)}(x) - \bigcap_{x\in\theta_\tau} \phi_{(\tau,\varphi)}(x)$

A vague class A is a class decided by a sample set θ_τ with subject τ and distinguish condition φ, which has two parts: *core region* and *border area*.

If a distinguish condition φ and a sample set θ_τ are given, and x is an object under subject τ, we then say that x is τ under condition φ and sample set θ_τ, for short, x is τ.

Based on Definition 2, we get:

FACT 1.
x is τ, if $x \in A(\theta_\tau, \varphi)$.
x is a standard τ, if $x \in \kappa(A(\theta_\tau, \varphi))$. If x is standard then x is not vague. All samples are standard.
x is a vague τ, if $x \in \mu(A(\theta_\tau, \varphi))$.
x is a sample of τ, if $x \in \theta_\tau$.

2.3 Vague degrees

Based on vague classes, we now discuss the concept of a vague degree. There are two kinds of vague degree: vague degrees of vague classes and vague degrees of objects. We start with the first.

When $A(\theta_\tau, \varphi)$ is finite, it may have a vague degree.

Definition 3 The vague degree of a vague class $A(\theta_\tau, \varphi)$ is denoted as $\mathrm{Vd}(A(\theta_\tau, \varphi))$

$$\mathrm{Vd}(A(\theta_\tau, \varphi)) = 1 - \frac{\|\kappa(A(\theta_\tau, \varphi))\|}{\|A(\theta_\tau, \varphi)\|}$$

The greater the value of $\kappa(A(\theta_\tau, \varphi))$, the smaller the vague index of $A(\theta_\tau, \varphi)$, and the more $A(\theta_\tau, \varphi)$ approaches to a precise class.

When $\|\kappa(A(\theta_\tau, \varphi))\| = \|A(\theta_\tau, \varphi)\|$, that is to say, every individual in $\|A(\theta_\tau, \varphi)\|$ is standard, we have $\mathrm{Vd}(A(\theta_\tau, \varphi)) = 0$.

Clear classes can be defined by means of vague degrees of vague classes:

Definition 4 $A(\theta_\tau, \varphi)$ is a clear class if and only if $\mathrm{Vd}(A(\theta_\tau, \varphi)) = 0$.

Clear class are a special case of vague classes. Normal sets are clear classes.

Proposition 1 $A(\theta_\tau, \varphi)$ is a clear class if and only if $\|\mu(A(\theta_\tau, \varphi)\| = 0$. In particular, $A(\theta_\tau, \varphi)$ does not have a border area.

Thus, the reverse of a vague class is a clear class. If $A(\theta_\tau, \varphi)$ has no border area, it is a clear class.

Now, we discuss the vague degree of an object x according to a vague class $A(\theta_\tau, \varphi)$.

Positive vague degree: The *positive vague degree* says to what extent an object x is τ. This is proportional to how many samples x are indistinguishable from. Intuitively, if x is indistinguishable from all samples, then "x is τ" is not vague. (If one is indistinguishable from all samples of "bald", then the vague degree of "he is bald" is 0: one is bald for sure, which is not vague at all.)

By definition 1, we obtain:

FACT 2. Let y be a sample ($y \in \theta_\tau$). x is indistinguishable with y if and only if $x \in \phi_{(\tau,\varphi)} y$.

Definition 5 $\mathrm{vd}^+(x, A(\theta_\tau, \varphi)) = 1 - \dfrac{\|\{\phi_{(\tau,\varphi)} y : x \in \phi_{(\tau,\varphi)} y, y \in \theta_\tau\}\|}{\|\theta_\tau\|}$

$\mathrm{vd}^+(x, A(\theta_\tau, \varphi))$ is the vague degree of 'x is τ'. If x is indistinguishable with all samples, then $\|\{\phi_{(\tau,\varphi)}y : x \in \phi_{(\tau,\varphi)}y, y \in \theta_\tau\}\| = \|\theta_\tau\|$, $\mathrm{vd}^+(x, A(\theta_\tau, \varphi)) = 0$, this means that x is a standard τ, and belongs to $A(\theta_\tau, \varphi)$ with no vagueness. Otherwise, we can get a number a, $0 < a \le 1$, the vague degree of 'x belongs to $A(\theta_\tau, \varphi)$' is a.

Negative vague degree: to what extent an object x is not τ.

To what extent an object x is not τ, it is proportional to how many samples x is distinguishable from. Intuitively, if x is distinguishable from all samples, then 'x is not τ' is not vague. (If one is different from all samples of bald, he is not bald for sure, then the vague degree of 'he is not bald' is 0.)

Definition 6 $\quad \mathrm{vd}^-(x, A(\theta_\tau, \varphi)) = 1 - \dfrac{\|\{\phi_{(\tau,\varphi)}y : x \notin \phi_{(\tau,\varphi)}y, y \in \theta_\tau\}\|}{\|\theta_\tau\|}$

When we make a judgement on whether an individual is τ, the greater the sample set is (that is, $\|\theta_\tau\|$ is greater), the more precise the vague degree is (whether positive or negative, it will acquire more decimal place). This indicates that, to get more precise vague degree, usually we need to find more samples.

3 Vague classes and a resolution of the Bald Paradox

3.1 The analysis on bald paradox

How do we judge whether a man is bald or not?

Whether an individual is bald or not is decided by an agent's visual impression. At the same time, being bald or not has a relation with the number of hairs. To judge whether an individual is bald, the number of hairs has some kind of power to decision power: up to a certain number, the individual must be bald, over another certain number, the individual cannot be bald.

Impressions are holistic and vague, while the number of hairs is discrete and can be accurately measured. So, there are two different perspectives: bald is a global 'macroscopic' impression, number (of hairs) is a result of close or 'microscopic' observation and count. In other words, impression is a cognitive outcome obtained in macroscopic perspective; the number of hairs is a cognitive outcome acquired by a cognitive measurement perspective. We will call these two perspectives whole or impression perspective versus measure or quantitative perspective, respectively.

Based on the above analysis, our views are these: the Bald Paradox is an epistemic problem, it should be discussed in an epistemic frame. It originates from the difference of two perspectives: impression and measure perspective. Repeated conversions between the two perspectives lead to puzzles.

Switching back and forth: *If a man with n hairs on his head is bald then a man with (n+1) hairs on his head is bald*. 'n hairs', '(n+1) hairs' come from a measure perspective, which focuses on hairs and sum of hairs. '. . . is bald' is an impression perspective, which focuses on the phenomenon of baldness.

If a man with n (number) hairs on his head is bald (impression) then a man with (n+1) (measure) hairs on his head is bald (impression).

We judge whether one is bald or not by a macroscopic perspective, not by counting the number of hairs. If we only use a macroscopic perspective, vagueness is possible, but there is no paradox. When measure perspective is added to impression, and the two perspectives switch to each other repeatedly, the paradox arises.

To resolve the paradox, we need to get rid of the connections or break the switching. It is not reasonable to get rid of the connections, because the number of hairs has a basic relation with being bald. So we consider breaking the switch. Suppose the upper bound is m, then, when counting the hairs to m, switching should be stop. Now the problem is how to set the upper bound, and to us, it is based on the concept of a vague class.

3.2 Vague classes and a resolution of the Bald Paradox

Definition 7 Let $f(x)$ be the sum of x' s hairs. $\tau = B$ (bald), For any $x \in \kappa(A(\theta_B, \varphi))$, we set

$G(A(\theta_B, \varphi)) = Max(\{f(x) : x \in \kappa(A(\theta_B, \varphi))\}$

$H(A(\theta_B, \varphi)) = Max(\{f(x) : x \in A(\theta_B, \varphi))\}$

$G(A(\theta_B, \varphi))$ and $H(A(\theta_B, \varphi))$ are two boundaries around $A(\theta_B, \varphi)$ from a hair sum viewpoint. $G(A(\theta_B, \varphi))$ is the upper bound of $A(\theta_B, \varphi)$'s core. $H(A(\theta_B, \varphi))$ is the upper bound of $A(\theta_B, \varphi)$.

We make two quick remarks:

If $f(a) \geq G(A(\theta_B, \varphi))$, that is, $x \notin \kappa(A(\theta_B, \varphi))$, a cannot be a standard bald object.

If $f(a) \geq H(A(\theta_B, \varphi))$, that is, $x \notin A(\theta_B, \varphi)$, a cannot be bald with sample set θ_B under condition φ.

Something should be underlined here. First we get a bald class by sense and impression, after that, we measure this class and get two features of this class: $G(A(\theta_B, \varphi))$ and $H(A(\theta_B, \varphi))$. It is not true, though, that we get or define bald classes by both $G(A(\theta_B, \varphi))$ and $H(A(\theta_B, \varphi))$.

For any $A(\theta_B, \varphi)$, due to the upper bound on the hair sum, the induction on hair sums in the Bald Paradox does not go through.

Instead, the inductive argument should be rewritten to a conditional induction:

Condition 1: for any $x, y \in A(\theta_B, \varphi)$, $f(y) = x + 1$, $f(y) < G(A(\theta_B, \varphi))$.

Condition 2: for any $x, y \in A(\theta_B, \varphi)$, $f(y) = x + 1$, $f(y) < H(A(\theta_B, \varphi))$.

When condition 1 is satisfied, if x is bald, then $y(f(y) = x + 1)$ is bald, and it is standard bald.

When condition 2 is satisfied, if x is bald, then $y(f(y) = x + 1)$ is bald, and y also could be non-standard bald.

The restrictive conditions (resulting in the upper bound) ensure that we will not introduce new individuals by this induction. Thus, the paradox is dispelled.

3.3 Answers to some possible questions

In summary, the basic method to dispel the Bald Paradox is this. According to intuition, the induction should have an upper bound, if it goes beyond the upper bound, the induction is not valid any more.

This method seems to be right, but there are several questions we need to answer.

Suppose there is an upper bound n,

(1) Why is n so exact? How to determine this n (the problem of selecting the upper bound)?

(2) Since n is an upper bound, and for bald, $n+1$ and n look no different, $n+1$ (or $n-1$) could also be upper bounds. Then, any n could be an upper bound. (The lifting of upper bound).

Here is our answer to these problems. n is determined by $A(\theta_B, \varphi)$, $A(\theta_B, \varphi)$ is determined by samples and distinguish power, n is not determined by information on the number of hairs. n is obtained by impression, not by measure. Under some conditions and impressions, we get n. We can also get $n + 1$ under other conditions and impressions, but that is because we are under different conditions. Under certain conditions, by impression we get only a certain n.

4 A formulation for vague classes

4.1 Language

Let $\mathcal{L}$ be a first order language, $\mathcal{L}^*$ is an extension of $\mathcal{L}$.

In $\mathcal{L}^*$, we add countably many new monadic predicate symbols Π (vague predicate) to $\mathcal{L}$, for any $P \in \Pi$, there are equality symbols $\approxeq_P$ and monadic predicates Θ_P. We add two types of formulas to the first-order formulas of $\mathcal{L}$: $x \approxeq_P y$ and $\Theta_P x$.

Definition 8: For any $P \in \Pi$,

$Px := \exists y(\Theta_P y \wedge x \approxeq_P y)$

$\odot Px := \forall y(\Theta_P y \to x \approxeq_P y)$

$\bigcirc Px := \exists y(\Theta_P y \wedge x \approxeq_P y) \wedge \exists y(\Theta_P y \wedge \neg x \approxeq_P y)$

Here are the intuitive meanings for these formulas:

$x \approxeq_P y$: x is indistinguishable from y on P[3]

$\Theta_P x$: x is a sample of P.

$\odot Px$: x is a standard P.

$\bigcirc Px$: x is a vague P.

Px: x is P.

Now we will show how to express various baldness assertions in $\mathcal{L}^*$.

[3]Whether 2 individuals are distinguishable is a problem that depends on an agent's cognitive ability. To make our basic analysis simpler, we did not introduce explicit parameters for agents. Then the language and semantics can be seen as characterizing one single agent's vague classes.

$\Theta_B x$: x is a sample of bald.
$Bx \coloneqq \exists y(\Theta_B y \wedge x \approxeq_B y)$
Bx: x is bald.
$\neg Bx$: x is not bald.
$\neg Bx \leftrightarrow \neg\exists y(\Theta_B y \wedge x \approxeq_B y) \leftrightarrow \forall y(\neg\Theta_B y \vee \neg(x \approxeq_B y)) \leftrightarrow \forall y(\Theta_B y \rightarrow \neg(x \approxeq_B y))$
$\odot Bx \coloneqq Bx \wedge \forall y(\Theta_B y \rightarrow x \approxeq_B y)$
$\odot Bx$: x is standard bald.
$\neg \odot Bx$: x is not standard bald.
$\neg \odot Bx \leftrightarrow \neg(Bx \wedge \forall y(\Theta_B y \rightarrow x \approxeq_B y))$
$\quad \leftrightarrow \neg Bx \vee \exists y \neg(\Theta_B y \rightarrow x \approxeq_B y) \leftrightarrow \neg Bx \vee \exists y(\Theta_B y \wedge \neg x \approxeq_B y)$
$Bx \wedge \neg \odot Bx$:x is non-standard bald, x is bald, but not standard bald.
$Bx \wedge \neg \odot Bx \leftrightarrow Bx \wedge \exists y(\Theta_B y \wedge \neg x \approxeq_B y)$
$\bigcirc Bx$: x is vague bald. (or x is bald vaguely)
$\bigcirc Bx \coloneqq Bx \wedge \exists y(\Theta_B y \wedge \neg x \approxeq_B y)$
$\bigcirc Bx \leftrightarrow Bx \wedge \neg \odot Bx$, A vague bald man is a non-standard bald man. Vice versa, a non-standard bald man is a vague bald man.
$\neg\bigcirc Bx$: x is not vague bald. (x is standard bald, or, x is not bald)
$\neg\bigcirc Bx \leftrightarrow \forall y(\Theta_B y \rightarrow x \approxeq_B y) \vee \forall y(\Theta_B y \rightarrow \neg(x \approxeq_B y))$
Bx, $\odot Bx$ and $\bigcirc Bx$ are abbreviations of complex formulas. Thus, all kinds of bald ($\odot Bx$, $\bigcirc Bx$ and even Bx, etc.) are not primitive notions, they are defined by samples and distinguishing power.

By definition 8, we can get:
$\bigcirc Bx \rightarrow Bx$, A vague bald man is a bald man.
$\odot Bx \rightarrow Bx$, A standard bald man is a bald man.
$\odot Bx \rightarrow \neg\bigcirc Bx$, A standard bald man is not a vague bald man (standard is not vague).
$\bigcirc Bx \rightarrow \neg \odot Bx$, A vague bald man is not a standard bald man (vague is not standard).
$\bigcirc Bx \leftrightarrow Bx \wedge \neg \odot Bx$, vague bald is equivalent with non-standard bald.

4.2 Model

The semantics for $\mathcal{L}^*$ extends the standard semantics for $\mathcal{L}$, composed of structure $\mathbb{A} = \langle D, \eta\rangle$, a valuation ρ and an interpretation $\sigma = \langle \mathbb{A}, \rho\rangle$. An interpretation is our notion of 'model'.

In $\mathcal{L}^*$, we add vague predicates set Π, and for any $P \in \Pi$, an equality symbol $\approxeq_P$ and monadic predicate Θ_P. In $\mathcal{L}^*$-semantic, we need some new elements to interpret these new symbols and the resulting new formulas.

Definition 9 A $\mathcal{L}^*$-structure is a tuple $\mathbb{B} = \langle D, \eta, T, \{\approx_\tau : \tau \in T\}, \theta, \xi\rangle$, a valuation ρ and an interpretation $\sigma = \langle \mathbb{B}, \rho\rangle$.

$\mathbb{B}$ is an expansion of $\mathbb{A} = \langle D, \eta\rangle$, adding the following components:

$T \neq \varnothing$ is a subject set (subject to distinguishing).

For any $\tau \in T$, $\approx_\tau$ is a reflexive and symmetry relation on D.

For any $\tau \in T$, θ is a sample choice function ($\theta : T \to \wp(D)$). For every subject τ, there is a sample set $\theta(\tau) \subseteq D$ with $\theta(\tau) \times \theta(\tau) \subseteq\approx_\tau$.

ξ: $\Pi \to T$. For any $P \in \Pi$, we interpret predicate P as a subject τ, or, assign a subject to every predicate, $\xi(P) = \tau$, ξ is similar to η.

Definition 10

(1) If α is a first-order $\mathcal{L}$-formula, then $\sigma(\alpha)$ is the same as in the definition for $\mathcal{L}$-semantic.

(2) $(\Theta_P(x))^\sigma = \begin{cases} 1, & if x^\sigma \in \theta_{\xi(P)} \\ 0, & otherwise \end{cases}$

(3) $x \approxeq_P y = \begin{cases} 1, & if x^\sigma \approx_{\xi(P)} y^\sigma \\ 0, & otherwise \end{cases}$

Definition 11 (Satisfiability and validity) Given a $\mathcal{L}^*$ -formula α. If there is an interpretation $\sigma = \langle V, \rho \rangle$, such that $\sigma(\alpha) = 1$, then α is satisfiable. If for any σ, there is $\sigma(\alpha) = 1$, then α is valid, denoted as $\vDash \alpha$.

Propositions 2

(1) $\vDash \bigcirc Bx \to Bx$

(2) $\vDash \odot Bx \to Bx$

(3) $\vDash \odot Bx \to \neg \bigcirc Bx$

(4) $\vDash \bigcirc Bx \to \neg \odot Bx$

(5) $\vDash \bigcirc Bx \leftrightarrow Bx \wedge \neg \odot Bx$

(6) $\vDash \exists y \Theta_B y \to (\odot Bx \to Bx)$

The intuitions for clauses (1) - (5) were given in Section 4.1. The meaning of (6) is this: under the condition that a standard bald man exists, a standard vague bald man is a bald man. The proofs of (1) through (6) are routine.

4.3 Some explanations on truth value

When we make a judgement on whether a man is bald, samples and the indistinguishability relation are very important factors. When we choose samples, it is easy and not vague, but when we judge whether a man is bald or not from samples, vagueness arises. Our definition of baldness does not deal with vagueness in judgement, it only says that, when the judgement is done, we get a bald class. The dividing line between bald and non-bald is strict, and bald and not bald are contradictory to each other. Thus, from the view point of truth values, our treatment of these predicates is 2-valued. In the same way, standard bald and non-standard bald are 2-valued too.

Intuitively, bald and non-bald should be contradictory relations. We cannot say that an individual is both bald and non-bald. On the other hand, sometimes, it is difficult to distinguish bald and non-bald strictly, there is a vague area, which is also according to intuition. What we can say is that the former intuition comes from logic, the latter intuition comes from cognitive reasons. In this paper, what we discuss is the concept bald, in other words, we discuss how to define bald. Under the 2-valued definition of bald, the dividing line between bald and non-bald is clear, not vague. This is a theoretical definition. Our theory does not focus on how to decide whether an individual is bald or not based on samples, which is a cognitive problem. In a practical judgement, the boundary between bald and non-bald is still vague, and relies on our impressions. That is our cognitive reality. Vague class theory does not solve the cognitive problems for assigning bald.

5 Comparison with other approaches

5.1 Impressive versus measure perspective

To capture the nature of vague expressions, this paper introduced the concept of a vague class. As we have mentioned in Section 2.1, in fuzzy theories, there are fuzzy sets corresponding to vague expressions. The core concept for a fuzzy set is degree of membership, which depends on what we called a scientific measure perspective. In our opinion, to understand a vague expression, what we use is an impressive perspective, not a measure perspective. Based on this view, we introduced vague classes, an impressive perspective-based concept.

5.2 Indistinguishability versus indifference relations

In this paper, a vague class has two basic elements: the *sample* and the *indistinguishable relation*. In our extended first-order language $\mathcal{L}^*$ with equality symbols $\approx_P$ for indistinguishability, and monadic predicates Θ_P for samples, we defined:

Bald: $Px := \exists y(\Theta_P y \wedge x \approx_P y)$

Standard bald: $\odot Px := \forall y(\Theta_P y \to x \approx_P y)$

Vague bald: $\bigcirc Px := \exists y(\Theta_P y \wedge x \approx_P y) \wedge \exists y(\Theta_P y \wedge \neg(x \approx_P y))$

Here is another approach. In [1], a representative paper on the *tolerant approach* to vagueness, the concept of an *indifference relation* is introduced. The principle of tolerance can then be stated as:

$\forall x \forall y(P(x) \wedge x \backsim_P y \to P(y))$

Here $\backsim_P$ is an indifference relation which is reflexive and symmetric, but possibly non-transitive. In the opinion of the authors, if the semantics of vague predicate is made sensitive to such indifference relations, the tolerate principle can be validate. Intuitive said, x is tall *tolerantly* if there exists an individual y such that x is similar to y by way of how tall x looks, and y is tall *classically*. In the specific paper mentioned, the authors defined three distinct notions of truth to capture vague predicates: *tolerant truth* (t-truth), *classical true* (c-truth), *strict true* (s-truth). The relations between t-truth and c-truth, s-truth and c-truth can be represented as follows[4]:

[4]For more details, see the cited paper.

$M \vDash^t P(\underline{a})$ iff $\exists d \backsim_P a : M \vDash^c P(\underline{d})$
$M \vDash^s P(\underline{a})$ iff $\exists d \backsim_P a : M \vDash^c P(\underline{d})$.

Now, we compare the two theories. At first glance, they are very similar in their intuitive ideas, and we even could find a rough correspondence, using the authors' notions of truth and our Definition 8 above:

Indistinguishable relation VS Indifference relation,
$\bigcirc Px$(vague bald) is true VS Px is t-true[5],
$\odot Px$ (standard bald) is true VS Px is s-true,
$\Theta_P y$(a sample of bald) is true VS Px is c-true.

But when we go into details, the differences appear.

Firstly, in this paper, we define Px, $\bigcirc Px$ and $\odot Px$ in an object language $\mathcal{L}^*$, here as the tolerance approach uses three kinds of truth, and tries to capture their distinctions through model theory. Second, to definePx, $\odot Px$, and $\bigcirc Px$, crucially for us, samples $\Theta_P y$ are introduced in this paper, which are based on an individual's cognition. In the cited tolerance approach, the part corresponding to $\Theta_P y$ is Py which is classically true, there is no cognitive perspective. Finally, our approach has only one truth, making the system 2-valued, whereas the tolerance approach uses at least three truth values.

5.3 Related work

What are the truth values of borderline cases? Different approaches have different ideas. In fact, many approaches try to capture vagueness by changing classical valuations. Three-valued approaches introduce a new truth value I (indefinite/indeterminate) to capture the intermediate state in between T (true) and F (false). In fuzzy theories, sentences are true to some degree in [0, 1][6]: for example, 'Tom is tall' may have a degree of 0.87. Supervaluationism introduces the technique of truth value gaps to vagueness. Fine [2] in particular proposed to account for the logical behavior of vague predicates by making a distinction between truth and 'supertruth'. A proposition is supertrue at a partial specification if it is classically true at all complete extensions. Epistemicists believe that there are strict boundaries between true and false. We cannot know where it is only because of agents' cognitive limitations: therefore, in epistemicism, classical 2-valuedness is preserved.

Let us briefly mention a few issues that face theories in the area. In technical aspects, many-valued approaches (including three-valued approaches like that of [1], fuzzy theories, and others, have many problems, both in validities [7] and in the

[5]The 'true' in left part is given in definition 8 of this paper, the 't-true', 'c-true', 's-true' are concepts of [1]

[6]For fuzzy theories, tautologies are true to degree 1, contradictions are true to degree 0, and contingent propositions may be true to any degree corresponding to a real number between 0 and 1.

[7]$P(a) \vee \neg P(a)$ is not a tautology, and $P(a) \wedge \neg P(a)$ is not a contradiction.

famous problem of 'penumbral connections'. [8] Moreover, higher-order vagueness is another prominent problem which many-valued approaches have to face. [9] The supervaluation approach can give a good analysis of penumbral connections, but it has problems with interpreting higher-order vagueness. Since 2-valued classical logic is preserved, epistemicism has no pressure when facing the technical problems mentioned above, but in a logical perspective, keeping 2-valued classical logic means that there is nothing new. If what we want to capture is how ordinary people (with imperfect logic) use vague expressions in daily life, classical 2-valued logic is not enough.

In this paper, there is a 2-valued definition of bald. From the perspective of technique, it is simpler than many-valued approaches and the supervaluation approach. On the other hand, different from the classical 2-valuedness of epistemicism, in this paper, vague predicates are defined based on samples and distinguishing power, and ordinary people's cognitive factors are considered and partly captured in our theory.

In addition to technical considerations, there is a deeper motive. We think that vague classes are based on human cognition, they do not primarily depend on truth. The concepts of truth and scientific measurement are acquired later on. To emphasize these ideas, we introduced the notion of sample $\Theta_P y$ which reflects our cognition, and we do not continue the traditional ways of emphasizing truth-value changes.

5.4 Sorites paradox

We add a few comments on solutions to the sorites paradox. The three valued approach ([13]) thinks that, because there is at least one not true (F or I) premise, we cannot get the conclusion by the standard inferences in the paradox. Fuzzy theories have two ways of dealing with the paradox. If the consequence relation requires keeping perfect truth, because the inductive premises have value almost 1 (but not perfectly true), we cannot get the conclusion. If the consequence relation itself requires maintaining true to a degree, then '→' is non-transitive in the analysis. In the supervaluation approach, the inductive premise is (super-)false under the interpretation, and the Sorites paradox is blocked. Epistemicists think that the inductive premises are not always true, and then the conclusion need not be true.

In short, there are mainly two ways to resolve the Sorites paradox: showing that

[8] As an illustration, consider the three-valued approach. (1) If a is indeterminate pink (P), red (R) and small (S), then $P(a) \wedge S(a)$ is indeed still indeterminate, but $P(a) \wedge R(a)$ should be false. Similarly with $P(a) \vee S(a)$ versus $P(a) \vee R(a)$: the former is indeterminate as desired, but we want the latter to be true. (2) Suppose that a and b are both borderline cases of tall individuals, though b is, in fact, a bit taller. Then $T(a) \wedge \neg T(b)$ should be false, but is predicated to be indeterminate. Fuzzy theories make similar wrong predictions for many other complex sentences as well.

[9] If we do not believe that there is a strict border between tall and not tall, then it is hard to believe that there are two strict borders between (i) the tall and the borderline cases, and (ii) the borderline cases and the not tall ones.

the inductive premise is not true, or cutting off the inference chains by denying transitivity. Every theory mentioned can resolve the paradox, but the key point is what they are based on (their interpretation for vagueness), and this should be able to make clear the mechanism underlying the paradox: why it is wrong, even though it looks reasonable.

In this paper, we give an analysis of the Bald Paradox, we judge whether one is bald by a macroscopic perspective, not by focusing on measuring numbers of hairs. If we only use a macroscopic perspective, vagueness is possible, but there is no paradox. When the measure perspective is added to that of impression, and the two perspectives switch to each other repeatedly, the paradox arises. The basic method to dispel the Bald Paradox our view is this: according to intuition, induction should have an upper bound, and if it goes beyond the upper bound, the induction is not valid any more.

6 Conclusion

In this paper, taking the simplest vague predicate "bald" as an example, we introduced the concept of a vague class. Based on vague classes, we gave a resolution to the Bald Paradox, and explained why it arises.

Vague classes are expressed by vague predicates. Vague predicates can be defined in an extended first-order language. Based on samples and distinguishing power, we defined bald, standard bald and vague bald, as an attempt to interpret vagueness in formal language. Constructing a new kind of predicate in a formal language to capture the nature of vague predication is a distinctive feature of this method. It is different from the usual truth-value changing approaches. Another special feature is our distinction between the impressive perspective and the measure perspective. In daily life, we usually use the impressive perspective when employing vague predicates. The Bald Paradox is a kind of sophistry making use of repeated switching between impression and measurement perspectives. The resolution method in this paper stops this conversion at a certain stage.

Acknowledgements This research is supported by the Major Project of National Social Science Foundation of China (No. 12&ZD119) and The Project-sponsored by SRF for ROCS, SEM. We especially want to thank Johan van Benthem, Fenrong Liu for their very helpful comments and suggestions. We also thank Yanjing Wang for helpful exchange and comments at workshops. Further thanks go to various colleagues and audiences for their valuable feedback at conference and workshops held in Beijing and Shanghai.

BIBLIOGRAPHY

[1] Cobreros, P., Ripley, D., Egre, P., van Rooij, R.A.M..: 'Tolerant, classical, strict'. *Journal of Philosophical Logic*, 41, 347-385, 2012.

[2] Fine, K.: 'Vagueness, Truth and Logic', *Synthese* 30, pp. 265-300, 1975.

[3] Graff, D.: 'Shifting Sands: An Interest-Relative Theory of Vagueness', *Philosophical Topics* 28, pp. 45-81, 2000.
[4] Kamp, H.: 'Two Theories About Adjectives', in Keenan, E L (ed.) *Formal Semantics of Natural Language*, Cambridge: Cambridge University Press, 1975.
[5] Raffman, D.: 'Vagueness and Context-Relativity', *Philosophical Studies* 81, pp.175-92, 1996.
[6] Soames, S. *Understanding Truth*, New York: Oxford University Press,1999.
[7] Soames, S.: 'Precis of Understanding Truth and replies', *Philosophy and Phenomenological Research*, LXV: 429-452, 2002.
[8] Sorensen, R. A.: 'Ambiguity, Discretion and the Sorites', *Monist* 81, pp. 217-35, 1998.
[9] Sorensen, R. A.: *Vagueness and Contradiction*, Oxford: Oxford University Press, 2001.
[10] Stanley, J.: 'Context, interest-relativity, and the sorites', *Analysis*, 63: 269-80, 2003.
[11] Smith, N. J. J., *Vagueness and Degrees of Trut*h, Oxford: Oxford University Press, 2008.
[12] Tye, M: 'Supervaluationism and the Law of Excluded Middle', *Analysis*, 49/3: 141-143, 1989.
[13] Tye. M.: 'Vague objects', *Mind*(99):535-57, 1990.
[14] van Rooij, R.A.M.: 'Vagueness, tolerance and non-transitive entailment'. In *Reasoning under Vagueness*. Berlin: Springer. 2011.
[15] Williamson, T.: *Vagueness*, London: Routledge, 1994.
[16] Williamson, T.: 'Soames on Vagueness', *Philosophy and Phenomenological Research* 65, pp. 422-8, 2002.
[17] Crispin, W.: 'The Epistemic Conception of Vagueness', *Southern Journal of Philosophy*, 33 (Supplement): 133-159, 1995.
[18] Zadeh, L.: 'Fuzzy Sets', *Information and Control* 8, pp. 338-53, 1965.