Vagueness and Context

Hans Kamp and Mark Sainsbury

19. Juni 2012

1 Part II. First steps towards a formal account: Semantics and Logic via supervaluation.

Things to be done in PART II:

1. Presentation of a model theory for a language L of Predicate Logic with one 1-place vague predicate P
2. The semantics and Logic of Supervaluation
3. Piece-wise precisification
4. Formal versions of the Sorites
5. Accounting for the Sorites using Supervaluation Theory
6. Adding Determinateness
7. The Logic(s) of Total and Partial Semantics

1. Model Theory for a language L of Predicate Logic with one 1-place vague predicate P

- Let L_0 be a language of standard first order predicate logic.

  We assume that all the standard operators of predicate logic – $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\forall$, $\exists$ – are primitive operators in L_0.

  N.B. This is to allow for the possibility of revising the semantics of some of these operators later on in ways that deviate from the definitions that one uses to define some operators in terms of others.

  (E.g. the definitions one would use to define $\vee$, $\rightarrow$, $\leftrightarrow$ and $\forall$ in case we decide to restrict the set of primitive operators to $\neg$, $\wedge$ and $\exists$).

- The non-logical constants of L_0 include predicates, individual constants and functors.

- We assume the standard model-theoretic semantics for L_0.

  In order to be prepared for the move to partial semantics when we add the vague predicate P to L_0 we split all clauses of the truth definition into a truth and a falsity clause.

(1) Def.

1. An *extensional model for* L_0 is a structure $M = < D, I >$, where D (the *domain*, or *universe*, of M) is a non-empty set and I is the interpretation function of M. I is defined for the non-logical constants of L_0. If c is an individual constant, then $I(c) \in D$, if Q is an n-place predicate, then $I(Q) \subseteq D^n$, and if f is an n-place functor, then $I(f)$ is a function from D^n to D.

2. An *assignment for* M is a function which assigns an element of D to each variable.

3. Semantic values of terms and formulas: $[\tau]^{M,g}$, the *value of* term τ in M under assignment g and $[\phi]^{M,g}$, the *truth value of* formula ϕ in M under assignment g, are defined as usual.

$-[v_i]^{M,g} = g(v_i)$; $[c]^{M,g} = I(c)$; $[f(\alpha_1, .., \alpha_n)]^{M,g} = I(f)([\alpha_1]^{M,g}, .., [\alpha_n]^{M,g})$.

$-[Q(\alpha_1, .., \alpha_n)]^{M,g} = 1$ if $< [\alpha_1]^{M,g}, .., [\alpha_n]^{M,g} > \in I(Q)$;
$-[Q(\alpha_1, .., \alpha_n)]^{M,g} = 0$ if $< [\alpha_1]^{M,g}, .., [\alpha_n]^{M,g} > \notin I(Q)$;

$-[\neg\phi]^{M,g} = 1$ iff $[\phi]^{M,g} = 0$;
$-[\neg\phi]^{M,g} = 0$ iff $[\phi]^{M,g} = 1$;

$-[\phi\&\psi]^{M,g} = 1$ iff $[\phi]^{M,g} = 1$ and $[\psi]^{M,g} = 1$;
$-[\phi\&\psi]^{M,g} = 0$ iff $[\phi]^{M,g} = 0$ or $[\psi]^{M,g} = 0$;

$-[\phi \vee \psi]^{M,g} = 1$ iff $[\phi]^{M,g} = 1$ or $[\psi]^{M,g} = 1$;
$-[\phi \vee \psi]^{M,g} = 0$ iff $[\phi]^{M,g} = 0$ and $[\psi]^{M,g} = 0$;

$-[\phi \rightarrow \psi]^{M,g} = 1$ iff $[\phi]^{M,g} = 0$ or $[\psi]^{M,g} = 1$;
$-[\phi \rightarrow \psi]^{M,g} = 0$ iff $[\phi]^{M,g} = 1$ and $[\psi]^{M,g} = 0$;

$-[\phi \leftrightarrow \psi]^{M,g} = 1$ iff $([\phi]^{M,g} = 1$ and $[\psi]^{M,g} = 1)$ or $([\phi]^{M,g} = 0$ and $[\psi]^{M,g} = 0)$;
$-[\phi \leftrightarrow \psi]^{M,g} = 0$ iff $([\phi]^{M,g} = 1$ and $[\psi]^{M,g} = 0)$ or $([\phi]^{M,g} = 0$ and $[\psi]^{M,g} = 1)$;

$-[(\forall v_i)\phi]^{M,g} = 1$ iff for every $d \in D[\phi]^{M,g[d/v_i]} = 1$;
$-[(\forall v_i)\phi]^{M,g} = 0$ iff for some $d \in D[\phi]^{M,g[d/v_i]} = 0$;

$-[(\exists v_i)\phi]^{M,g} = 1$ iff for some $d \in D[\phi]^{M,g[d/v_i]} = 1$;
$-[(\exists v_i)\phi]^{M,g} = 0$ iff for every $d \in D[\phi]^{M,g[d/v_i]} = 0$.

4. If ϕ has no free variables, then $[\phi]^{M,g}$ does not depend on g, so we may write '$[\phi]^{M}$', omitting g.

5. Logical consequence and logical truth are also defined in the familiar way:

–The sentence ϕ is a *logical consequence of* the set of sentences Γ iff for every model M, if $[\psi]^M = 1$ for all $\psi \in \Gamma$, then $[\phi]^M = 1$.

–ϕ is a *logical truth* iff it is a logical consequence of the empty set of premises.

- Models for L are obtained from models for L_0 by adding an interpretation for P.

- A simple way to model the vagueness of P is to assume that its interpretations in models M for L consist of an *extension* $I^+_M(P)$ and an *anti-extension* $I^-_M(P)$.

  (Instead of 'extension' and 'anti-extension' one also often speaks of 'positive extension' and 'negative extension').

  The extension of P in M consists of the elements of D that are clear cases of P according to M and the anti-extension of P in M of the elements that are clearly not cases of P.

  But in addition to the extension and the anti-extesion there may be *borderline cases of P in M* – elements of D that are neither clear cases of P nor clear non-cases of P. When this is so, we say that 'P is vague according to M'.

We assume that it is always the case that $I^+(P) \cap I^-(P) = \emptyset$. But the vagueness of P according to M manifests itself in that $I^+(P) \cup I^-(P) \neq D$. We refer to $D - (I^+(P) \cap I^+(P) \neq D)$ as the *truth value gap of P in M*.

- Let M be a model for L. We can define truth in M in almost exactly the same way as we did for L_0. We only need an extra clause for atomic formulas of the form '$P(\alpha)$'. The standard pair of clauses is the following:

(2) $-[P(\alpha)]^{M,g} = 1$ if $[\alpha]^{M,g} \in I^+(P)$;
$-[[P(\alpha)]^{M,g} = 0$ if $[\alpha]^{M,g} \in I^-(P)$.

Note that the addition of this clause renders the truth definition partial: $[\phi]^{M,g}$ may de undefined.

- The partiality of the truth definition for L has consequences for the logic generated by the definition for logical consequence).

For instance, many formulas of the form '$\phi \vee \neg\phi$' do no longer come out as tautologies.

- The logic generated by the truth definition (1.3) supplemented with the clauses in (2) and the definition of logical consequence given by (1.5) is the so-called 'Strong Kleene Logic'.

- Does this establish that when vagueness comes into play, logic can no longer be classical? No, not automatically. The partial models for L we have just defined allow for other ways of defining the logical consequence relation.

- A number of such ways are made available by *supervaluation*.

2. The semantics and Logic of Supervaluation

Supervaluation is based on the following idea:

If P is vague in that it admits of borderline cases, its boderline cases could be resolved one way or another and P's vagueness thereby removed.

That is, in conjunction with each partial model M for L we can consider all possible ways in which the truth value gap of P in M can be closed. Each of these ways gives us a 'classical' model N for L, in which there is no truth value gap for P and in which $[\phi]^{N,g}$ is always defined.

- Such classical models for L, which ate like M except that all borderline cases of P in M have been resolved, are called *complete precisifications of M*.

- Complete precisifications of a model M for L are a special case of models for L that are *sharpenings* of M.

  In general, a model $N = < D', I' >$ for L is a *sharpening* of a model $M =< D, I >$ (in symbols: $M \preceq N$) iff (a) $D' = D$; (b) for every non-logical constant β of L_0, $I'(\beta) = I(\beta)$ and
  (c) $I^+(P) \subseteq I'^+(P)$ and $I^-(P) \subseteq I'^-(P)$.

- The truth definition for L is monotonic: truth and falsity are preserved by sharpening:

  When $M \preceq N$, then for any formula ϕ and assignment g, if $[\phi]^{M,g} = 1/0$, then $[\phi]^{N,g} = 1/0$.

- A pair $< M, \mathcal{N} >$, where $\mathcal{N}$ is a set of complete precisifications of M, is called a *supermodel* for L.

  M is called the *base model* of $< M, \mathcal{N} >$ and the members of $\mathcal{N}$ the (*complete*) *precisifications* of $< M, \mathcal{N} >$.

- In a supermodel $\mathbf{M} =< M, \mathcal{N} >$ 'truth values' can be defined in more than one way.

  i. We can define them just as before, looking only at M and ignoring $\mathcal{N}$.

  ii. We can define a sentence ϕ as *supertrue* in $\mathbf{M}$ iff $[\phi]^N = 1$ for every $N \in \mathcal{N}$; and, likewise, ϕ as *superfalse* in $\mathbf{M}$ iff $[\phi]^N = 0$ for every $N \in \mathcal{N}$.

- Like truth as defined directly on M, supertruth is in general a partial notion.

  In particular, when an atomic sentence $P(c)$ is without a truth value in M, it may be expected to also lack a supertruth-value.

  $P(c)$ is bound to lack a supertruth-value in $N \in \mathcal{N}$ if $\mathcal{N}$ contains all formally possible complete specifications of M.

- However, all formulas that are theorems of classical logic come out as supertrue. In that sense supertruth preserves classical logic.

- We now have new options for defining logical consequence (besides the definition we already gave, which only refers to M). Here are two such options:

(3) a. *Global Logical Consequence*

The sentence ϕ is a *global logical consequence*$_{sg}$ *of* the set of sentences Γ iff for every supermodel $\mathbf{M}$ if $[\psi]^M$ is supertrue in $\mathbf{M}$ for all $\psi \in \Gamma$,then $[\phi]^M$ is supertrue in $\mathbf{M}$.

b. *Local Logical Consequence*

The sentence ϕ is a *local logical consequence*$_{sl}$ *of* the set of sentences Γ iff for every supermodel $< M, \mathcal{N} >$, and every $N \in \mathcal{N}$: if $[\psi]^N = 1$ for all $\psi \in \Gamma$,then $[\psi]^N = 1$.

- It is not hard to see that these two consequence relations generate the same logic, viz. classical logic. But conceptually the two notions are quite different.

- The second of these definitions reflects the following intuition:

  (4) *Intuitive Justification of Local Logical Consequence*

  > It may be indeterminate for some individuals whether they satisfy the predicate P that occurs in the premise(s) and conclusion of my argument. But what I am really interested in is whether the truth of the conclusion is guaranteed by the truth of the premises when such cases are resolved, no matter how they are resolved, as long as they are resolved in the same way in premises and conclusion.

- Global Logical Consequence will be of interest to someone who thinks that where vagueness is present supertruth is the natural notion of truth, and that validity should therefore amount to preservation of supertruth.

  If you do not share the intuition that supertruth – truth no matter how borderline cases are resolved – is the right notion of truth in the presence of vagueness, then Global Logical Consequence will not have much intuitive appeal either.

- **Penumbral Connections**

- What can we say about *which* complete precisifications should be part of a supermodel $< M, \mathcal{N} >$? That depends on our perspective.

  Perspective 1: The predicate P models the phenomenon of borderline cases in a purely abstract and general way: any way of resolving the truth value gap of P in a partial model M is admitted.

  From this perspective each partial model M determines a unique supermodel $< M, \mathcal{N} >$, that in which $\mathcal{N}$ is the set of all possible complete precisifications of M.

  Perspective 2: P is thought of as a proxy for any one of the actual vague predicates found in language or thought.

  These actual predicates may be subject to various constraints, which given some particular partial model M, rule out some of the possible complete precisifications of M.

  If we do not know about what these constraints could be, the best we can do is to admit as supermodels all structures $< M, \mathcal{N} >$ such that $\mathcal{N}$ is some non-empty set of complete precisifications of M.

Perspective 3: We use P to represent some particular vague predicate, such as *red*, *tall*, *cup*, *rain*,.... The semantics of such predicates is often subject to special constraints.

For instance, two color patches d and d' may be both borderline cases of 'red', but it may nevertheless be the case that d' is redder than d. This imposes a constraint on how these borderline cases may be resolved: any precisification that puts d into the extension of 'red' should put d' into its extension as well.

In such cases, where P represents some particular vague predicate, admissible supermodels $< M, \mathcal{N} >$ for a given partial model M will reflect the resolution constraints for that predicate: the members $N \in \mathcal{N}$ are all and only those that are compatible with these constraints.

- (Fine 1975) discusses penumbral connections between pairs of vague predicates.

  Example: the adjectives 'red' and 'pink': d and d' can be on the borderline between 'red' and 'pink, but d' may be redder than d. So complete precisifications N in which d is in the extension of 'red' while d' is in the extension of 'pink' are ruled out.

- N.B. A formal treatment of Fine's example requires a formal language with more than one vague predicate. Extending L to a language with several vague predicates is straightforward.

- For some natural language predicates the constraints they impose on the resolution of borderline cases are fairly easy to identify.

  Example: '1-dimensional gradable adjectives like 'tall' or 'heavy'.

  These are governed by the following *1-dimensionality constraint*:

  (5) (1-dimensionality constraint)

  > For any two objects d and d' within the application range of P either d satisfies P to at least the same degree as d' or d' satisfies P to at least the same degree as d.

In (Kamp 1975) supermodels satisfying constraints like the 1-dimensionality constraint are used to define the comparative forms of adjectives as involving an operator $Comp$ which turns adjectives into the 2-place predicates that are expressed by their comparatives:

For any 1-dimensional vague predicate P the semantics of $Comp(P)$ is given by:

(6) $[Comp(P)(x, y)]^{\mathbf{M},g} = 1$ if

(i) for all $N \in \mathcal{N}$, if $[P(y)]^{N,g} = 1$, then $[P(x)]^{N,g} = 1$ and

(ii) there is some $N \in \mathcal{N}$ such that $[P(x)]^{N,g} = 1$ and $[P(y)]^{N,g} = 0$

It can be shown that if P is 1-dimensional, then $Comp(P)$ defines a strict linear order.

- N.B. This way of analyzing comparatives has been disputed. Instead many linguistic analyses of gradable adjectives assume that they are not 1-place predicates but 2-place predicates, which express relations between entities and *degrees*.

  Another problem for the analysis of comparatives via 'Comp' is that comparative relations hold not only between elements from the borderline area but also between objects both of which are in the positive extension of P (or both in its anti-extension).

  To make (6) work for such objects we have to allow the semantics of P to vary so that in some models M these objects become borderline cases too.

  Perhaps such models can be thought of as reflecting special contexts in which the semantics of P deviates drastically from what it normally is.

3. Piece-wise precisification

- The idea of the complete precisification of a vague predicate is obviously an idealization.

  A more realistic view of how vague predicates can be sharpened is that often they are only sharpened only bit by bit.

  Complete specifications might still be thought of as virtual results, as limits of converging sequences of piece-wise sharpenings.

This brings us to the notion of a *Precisification Structure*:

(7) Def. A *Precisification Structure for* L is a pair $< M, \mathcal{N} >$, where:

(i) M is a partial model for L;
(ii) $\mathcal{N}$ is a partially ordered structure of (possibly incomplete) sharpenings of M.

(The partial ordering of $\mathcal{N}$ is the relation $\preceq$:
$N \preceq N'$ iff N' is a sharpening of N or $N' = N$.

We refer to M as the *base model* of $< M, \mathcal{N} >$.
It is convenient to assume that M itself is a member of $\mathcal{N}$.

(N.B. Shapiro refers to Precisification Structures as *frames*.)

- Even when a Precisification Structure contains no complete precisifications, it may generate a supermodel 'in the limit'.

  Let $< M, \mathcal{N} >$ be a Precisification Structure and let B be a maximal branch of $< M, \mathcal{N} >$. We can define the *limit* of B as follows:

  $I^+_{N_B}(P) = \cup\{I^+_N(P) \mid N \in B\}$; $I^-_{N_B}(P) = \cup\{I^-_N(P) \mid N \in B\}$; otherwise N_B is like all the members of $\mathcal{N}$.

  It is easy to see that N_B is a sharpening of all the models N in B and that it is the least precise such model.

  Note also that the limit of a branch can be a complete precisification even though none of the members of the branch are.

  (N.B. B is a branch of $\preceq$ iff it is a linearly ordered subset of $\preceq$. It is a maximal branch of $\preceq$ if it is a branch and there is no other branch of $\preceq$ which properly includes it.)

- A Precisification Structure $< M, \mathcal{N} >$ is said to *generate a supermodel* iff the limit of each of its branches is a complete precisification.

- Precisification Structures with this property can be used in the semantics of P in just the way we have been using supermodels. The only difference is in how the complete precisifications come about.

  (A proposal along these lines was made in (Kamp 1975).

  In many applications the notion of a supermodel is problematic, whether its complete precisifications are postulated or construed as limits.)

- Can Precisification Structures also play some other role in the semantics of P than as generators of supermodels?

  The first question we must answer in this connection is: How are we to think of what the members N of $\mathcal{N}$ stand for?

One way to think of the $N \in \mathcal{N}$ is as the results of contextual modification of the base model M.

This interpretation is closely tied to the view that context-dependence is a key feature of vagueness:

the phenomenon of borderline cases does not only imply the freedom of choice that individual users have to decide such cases, but also contextual malleability in a more general sense:

contextual adaptation can be brought about by choices that are made by speakers but also by contextual factors that need not be within their control.

- We will return to these questions at length.

For now just suppose that the different members of $\mathcal{N}$ correspond to different contexts of use for P.

Also suppose also that the relation $\preceq$ can be interpreted as a kind of 'accessibility relation' between contexts:

'$N \preceq N'$' means that a context reflected by N can develop into a context reflected by N'.

- This makes Precisification Structures look like *Kripke structures*:

  A *Kripke structure* is a pair $< W, R >$ in which W is a non-empty set (of 'worlds', or 'indices') and R is an accessibility relation between the members of W, and where the members of W determine the semantic values of the expressions of some language.

  Kripke structures have been extensively used in the analysis of a wide range of logical notions and linguistic phenonema.

  The formal similarity between Precisification Structures and Kripke structures makes it tempting to employ concepts and techniques from modal logic (in this very general sense of the term) also to the analysis of vagueness.

  We will return to this point.

- If the members of the set $\mathcal{N}$ of a precision structure are to be thought of as contexts, then an analysis along 'modal' lines is likely to make vagueness a semantic phenomenon that is context-dependent in two distinct ways:

  (a) the extensions of vague predicates can vary as a function of context;

  (b) the semantics of certain operators O of the language is context dependent in that the semantic value of an expression obtained by applying O in one context may depend on the values of the expressions to which O has been applied in other contexts.

  It is important to see the difference between these two kinds of context dependence. We cannot have (b) without (a). But we can have (a) without (b).

- One way in which theories of vagueness that assign an important role to context differ is just this:

  Does context only affect the extensions of certain lexical predicates, or does it also affect the semantics of logical operators? And if so, which operators?

Where we got so far:

1. **Presentation of a model theory for a language L of Predicate Logic with one 1-place vague predicate** P

2. **The semantics and Logic of Supervaluation**

3. **Piece-wise precisification**

4. Formal versions of the Sorites

5. Accounting for the Sorites using Supervaluation Theory

6. Adding Determinateness

7. The Logic(s) of Total and Partial Semantics

4. Formal versions of the Sorites

- One of the tasks of a semantics and logic for vague predicates is to account for the Sorites Paradox.

- It is widely held that such an account must accomplish two things:

1. It must show either that the paradoxical argument is, contrary to first impressions, unsound:

   either it is not valid (in that it involves the application of invalid logical principles) or else not all its premises are true.

2. It must give some kind of explanation of where the impression that it is a sound argument comes from.

- Formal presentations of the Sorites come in two forms.

  Both forms involve a conclusion of the form $P(a_N)$ and one premise of the form $P(a_0)$.

  Here 'a_0' and 'a_N' are individual constants which denote objects of which it is unequivocally true that the first does satisfy P, whereas the second does not.

  Furthermore, we assume that $a_1, \ldots, a_{N-1}$ are objects such that for n = 0, ..., N-1, a_n and a_{n+1} are within the tolerance margin of P.

  Two objects d and d' are *within the tolerance margin for P* iff the differences between d and d' (if any) are irrelevant to satisfaction of P.

  We express this relation formally as '$\approx_P$.

- The two versions differ with regard to the remaining premises.

  The first version has N additional premises, each of which is a conditional of the form '$P(a_n) \rightarrow P(a_{n+1})$' (again, for n = 0, …, N-1).

  The second version is usually given with one additional premise, which bundles the conditional premises of the first version into a single universally quantified conjunction.

  What we want for the second premise is something like this:

  '$(\forall n)(P(a_n) \rightarrow P(a_{n+1}))$'.

  But note that this is not a well-formed formula.

To express the universalised condition correctly we need a function f from objects to objects which maps any object d to an object $f(d)$ such that d and $f(d)$ are within the tolerance margin for P.

Moreover, there must be an object d and a natural number N such that $P(d)$ is true and $P(f^N(d))$ is false.

Given these properties of f we can state the paradoxical argument as in (8.b).

An alternative for the second form is given in (8.c).

This form makes use of the general tolerance principle in (8.d).
An additional third premise is necessary to state that arguments and values of f are always within their P-margin of tolerance.

(8) a. $P(a_0)$
$P(a_0) \rightarrow P(a_1)$
$P(a_1) \rightarrow P(a_2)$
$\vdots$
$P(a_{N-1}) \rightarrow P(a_N)$

Therefore:

$P(a_N)$

b. $P(a_0)$
$(\forall x)(P(x) \rightarrow P(f(x)))$

Therefore:

$P(f^N(a_0))$

c. $P(a_0)$
$(\forall x)(\forall y)(\approx_P (x, y) \rightarrow (P(x) \rightarrow P(y)))$
$(\forall x) \approx_P (x, f(x))$

Therefore:

$P(f^N(a_0))$

d. $(\forall x)(\forall y)(\approx_P (x, y) \rightarrow (P(x) \rightarrow P(y)))$

5. Accounting for the Sorites using Supervaluation Theory

- The semantics and logic of P in terms of supermodels and the local definition of logical consequence give us some tools to deal with the Sorites Paradox .

- We start with version (8.a).

  First, note that in any supermodel $< M, \mathcal{N} >$ such that $P(a_0)$ is true in M and $P(a_N)$ is false in M it must be the case that:

  for any complete precisification $N \in \mathcal{N}$ there is some n $\leq$ N-1 such that a_n is in $I_N(P)$ and a_{n+1} is not.

  This means that for each $N \in \mathcal{N}$ one of the conditional premises is false. So there is no $N \in \mathcal{N}$ in which the premises are all true. (And so the conjunction of the premises is superfalse in $< M, \mathcal{N} >$, just as the conclusion $P(a_N)$.)

  Since this holds for every supermodel, the conclusion $P(a_N)$ is a logical consequence of the premises, by default. But the argument cannot be sound.

- Second, the impression that the premises of (8.a) are true while the conclusion is false can be given some kind of explanation along the following lines.

  Consider again a supermodel $< M, \mathcal{N} >$ such that $P(a_0)$ is true and $P(a_N)$ false in M.

  Take any conditional premise '$P(a_n) \to P(a_{n+1})$'.

  Let $S(n, n+1)$ be the set of those $N \in \mathcal{N}$ such that a_n is in $I_N(P)$ and a_{n+1} is not.

  So $S(n, n+1)$ is the set of the N's in which the borderline between the P's and the non-P's runs between a_n and a_{n+1}.

But there are many, many other ways in which the borderline can be drawn, viz. between a_i and a_{i+1} for any i $\neq$ n. These other ways correspond to precisifications $N \in \mathcal{N}$ that are not in $S(n, n+1)$.

Surely this second set, $\mathcal{N} \setminus S(n, n+1)$, is much bigger than $S(n, n+1)$. So, loosely speaking, '$P(a_n) \rightarrow P(a_{n+1})$' is 'very nearly true' in $< M, \mathcal{N} >$, in that it is true in the vast majority of its complete precisifications.

- An explanation along these lines of why a Sorites argument in the form of (8.a) might strike us as valid was suggested many years ago by Rich Thomason.

  Thomason's suggestion was based on the supervaluation account of (Kamp 1975).

  It makes use of a feature of the models proposed there that we have not yet spoken about.

- (Kamp 1975) contains a critical discussion of Fuzzy Logic, pointing out that this approach has a fundamental problem with stating the semantics of the binary connectives (among them 'and' and 'or').

  As an approximation of what Fuzzy Logic seems to want, the paper offers a way of assigning 'credibility values' to sentences of L in supermodels $< M, \mathcal{N} >$ via a probability function over $\mathcal{N}$.

  For the supermodels under discussion a plausible probability function will assign only a small value to any of the sets $S(n, n+1)$ and a correspondingly high value to their complements $\mathcal{N} \setminus S(n, n+1)$.

  Such a probability measure assigns to all conditionals '$P(a_n) \to P(a_{n+1})$' a very high value: either 1 or else close to 1.

This explains why an argument of the form (8.a) may look like a (pretty) sound argument: it is valid and all its premises are either true or nearly true.

What this impression overlooks is although the premises are in a certain sense all fully or nearly true, this is not so for their conjunction.

- There is another aspect to these considerations, which is proof-theoretic rather than semantic:

  To obtain the conclusion of (8.a) from its premises all we need is a series of applications of the rule of Modus Ponens.

  What could be wrong with a series of applications of *that* (widely used and deeply trusted) inference rule? Can a string of such rule applications possibly lead from nearly true premises to an radically false conclusion?

  The answer is ‘ yes, it can!’

Within Fuzzy Logic, and also in other logics that employ some notion of ‘degree of truth’, M.P. is not strictly valid;

when M.P. is applied to two premises A and $A \rightarrow B$, neither of which is fully true, then the result B may have a value that is less than the values of either of the premises.

In this way successive applications of M.P. may lead from nearly true premises eventually to a conclusion that is as far from truth as you like.

- This is the essence of the accounts of the Sorites argument that have been given by advocates of Fuzzy Logic and other degree logics.

  To repeat: None of the premises of the argument is actually false. However, some of them are not fully true, but only close to true.

  But the relevant logic is not classical: classically valid inference rules, which are guaranteed to preserve full truth, do not necessarily preserve other degrees of truth. And that is what’s wrong with the argument in (8.a).

Within Fuzzy Logic you should not apply a rule like M.P. without checking how much you may have thereby lowered the 'fuzzy truth value (or 'degree of truth'). The illusion that the argument is valid comes from forgetting that there is this danger as soon as one departs from the 100% true.

- Note that this is a quite different way of dealing with the Sorites than the one involving supermodels and locally defined logical consequence.

  It rests on adopting a non-classical logic, rather than keeping classical logic intact, but arguing for the impossibility of simultaneously satisfying all premises.

- In view of our commitments to other aspects of vagueness in these lectures we will not pursue the degree.theoretical approaches any further.

Account of (8.b).

In any complete precisification N of a supermodel $< M, \mathcal{N} >$ in which $P(a_0)$ is true and $P(a_N)$ false the second, universally quantified premise is false.

For in order for it to be true in N, all its instances to constants a_n (with 0 $\leq$ n $\leq$ N-1) must be true.

But that is impossible: since N must verify $P(a_0)$ and falsify $P(a_N)$, it must draw a line between some a_n and its successor. So it will falsify '$P(a_n) \rightarrow P(a_{n+1})$', and with that univerdal quantification of which this is an instance.

- Second, it is easy to fall victim to the impression that the second premise should be true, or at least nearly true.

  That illusion comes from confusing the truth or near-truth of the instances of a universal generalization with the truth or near-truth of the universal generalization itself: even if its instances are all close to true, the generalization, which is equivalent to the conjunction of all its instances, may be false.

- Account of (8.c).

  Let us assume that the supermodel $< M, \mathcal{N} >$ verifies both $P(a_0)$ and $(\forall x) \approx_P (x, f(x))$ (in the strong sense that they are true in M); and let $P(f^N(a_0)$ be false in M.

  Then the second premise must be false in every $N \in \mathcal{N}$.

  For N must falsify at least one conditional of the form '$P(f^n(a_0) \rightarrow P(f^{n+1}(a_0)$' (with $0 \leq$ n $\leq$ N-1).

  And since '$\approx_P (f^n(a_0), f^{n+1}(a_0))$' is true in N (as an instance of the third premise, which is true in M and thus in each of its complete precisifications), the following instance of the second premise is false in N:

  '$\approx_P (f^n(a_0), f^{n+1}(a_0)) \rightarrow (P(f^n(a_0)) \rightarrow P(f^{n+1}(a_0)))$

Since this holds for all $N \in \mathcal{N}$, the second premise is superfalse in $< M, \mathcal{N} >$, just as we concluded above for the second premise of (8.b).

- The impression of the soundness of (8.c) is accounted for as before:

  the second premise (now the Tolerance Principle) strikes us as true or nearly true because we confuse the near-truth of the universal quantification with the near-truth of its instances.

- How good an account of the various forms of the Sorites Paradox is this?

  Judgments seem to differ on this score. Some are unconvinced because they feel that the conditional premises of (8.a) and the second premises of (8.b) and (8.c) really are in an important sense true; this isn't just an illusion!

  Aeguably the supervaluation story fails to do justice to that intuition.

- But note well: For someone who resists a solution to the Sorites according to which at least one premise of any Sorites argument is false there are only two options:

  (i) the *logic* used in arriving at the conclusion is at fault.

  As we have seen in connection with (8.a), this position entails at a minimum giving up Modus Ponens. Some other inference principles may be expected to go by the wayside too.

(ii) Sorites arguments can be sound, in the sense that (a) they are valid and yet (b) their conclusions can be false while their premises are true.

According to this view the relation between the world and the way we think and reason about it is truly paradoxical:

Valid principles of reasoning can lead from true premises to false conclusions; that's life and we just have to do the best we can not to stay clear of such traps.

But then, what can we do to avoid those traps?

6. Adding Determinateness

- All the models we have thus far discussed – partial models, supermodels and Precisification Structures – assign to P a semantic value consisting of three parts: extension, anti-extension and truth value gap.

  But as things stand, this tripartite division is not expressible within L itself.

  In particular, we cannot express that a is in the truth value gap of P. Just try!

  A first attempt might go like this:

'$P(a)$' expresses that a is in the extension of P, and '$\neg P(a)$' that a is in the anti-extension of P. So if a is in the truth value gap of P – that is, if neither of these two cases applies – then both of these formulas must be false, and thus their negations must be true:

(9) $\neg P(a)$ & $\neg\neg P(a)$

But of course this won't do as an expression of what we want. According to the semantics of $\neg$ we have given, (9) is just a plain contradiction.

What we need is a 'weak' negation $\sim$, such that $\sim \phi$ is true if ϕ is either false or undefined.

But so far our language L doesn't have such a negation. And adding such a negation to L has the disadvantage that it makes the semantics non-monotonic.

(For instance, if $N \prec N'$ and a is in the extension of P in N' but not in the extension of P in N, then $\sim P(a)$ will be true in N but false in N'.)

- Another way to obtain the expressive power needed to express that a is in the truth value gap of P is to add a 'determinateness operator' $\triangle$.

  $\triangle$ is a 1-place sentence operator, like $\neg$ and $\sim$, and it expresses that its operand is ('definitely') true. In other words:

  $\triangle\phi$ is true if ϕ is true and false if ϕ is either false or undefined.

- With the help of $\triangle$ we can express that a is in the truth value gap of P as in (10):

  (10) $\neg \triangle P(a)$ & $\neg \triangle \neg P(a)$

  From what has been said so far it would seem that $\triangle$ destroys monotonicity in just the same way that $\sim$ does.

  But there is a way of defining the semantics of $\triangle$ that preserves monotonicity:

  We say that $\triangle\phi$ is true at any N of our model structure iff ϕ is true in the base model M.

- In connection with supermodels this is all we can want: Both for the base model M and for all complete precisifications N we get the truth clause:

(11) a. $[\triangle\phi]^{M/N,g} = 1$ if $[\phi]^{M,g} = 1$
b. $[\triangle\phi]^{M/N,g} = 0$ if either $[\phi]^{M,g} = 0$ or $[\phi]^{M,g}$ is undefined

Note well:

(i) Definition (11) relies on the partial truth definition for L in the partial model M.

(ii) (11) has the arguably counterintuitive effect that in certain complete precisifications N $\triangle\phi$ is true and at the same time ϕ is false.

For those who see the role that complete precisifications of supermodels play in the characterization of truth and validity merely as a technical device, (ii) need not be a ground for worry.

- For Precisification Structures the matter is different.

  Let us assume again that the members N of $\mathcal{N}$ play the part of, or correspond to, different contexts of use.

  Then we may want to be able to express facts about the truth value gaps of P not only in M but also in these N's.

  The definition of $\triangle$ in (11) does not give us this. And an operator that does will inevitably lead to non-monotonicity.

- Note that from the formal perspective of modal logic $\triangle$ functions as a necessity operator satisfying the axioms of S5. This entails in particular, iteration of $\triangle$ is redundant.

- **Determinateness and Higher Order Vagueness**

  Let L' be the language L + ($\triangle$).

  According to the semantics for $\triangle$ given in (11) all formulas of L' that are of the form $\triangle\phi$ (as well as all logical compounds of such formulas) are bivalent.

  This is so in particular for all formulas of the form $\triangle P(x)$; and that means that the language L' treats the separation between extension and truth value gap as sharp:

  There are no a such that $\triangle P(\underline{a})$ is indeterminate (in the sense of lacking a definite truth value).

  (The separation between truth value gap and anti-extension is treated as sharp in the same sense.)

- From the earliest formal treatments of vagueness this is an aspect of the semantics that (11) assigns to $\triangle$ that has been seen as problematic.

- The problem is addressed explicitly in (Fine 1975). Fine shows how to formulate a model-theoretic semantics for L' in which the borderline between extension and truth value gap and that between truth value gap and anti-extension are also fuzzy.

  Furthermore, in his formal treatment the borderline between the extension and the first of these second order gaps can be fuzzy as well; and so on, and on.

- On the one hand such considerations may seem compelling: if there is a borderline area between extension and anti-extension to begin with, what plausibility is there in the assumption that there are sharp boundaries that separate this borderline area from extension and anti-extension?

  But on the other hand it is hard to see what kind of *positive* evidence can be adduced for higher order borderline areas.

  Our intuitions about borderline cases of a predicate P of our language (such as 'red' or 'bald' or 'heap') are often reasonably tangible:

  we are often in situations where we do not know whether some person a should be described as bald, and where we feel that our not knowing what to say isn't because of a lack of actual information about a, or because we do not know enough about the meaning of 'bald'. In such a situation we are inclined to say that a is a borderline case of 'bald' (or, put slightly differently, that 'a is in the truth value gap of 'bald'').

  But can we make out a meaningful distinction between being confident in classifying a as a borderline case of 'bald' and hesitating over whether a should be classified as a borderline case of baldness or as a definite case of baldness?

- Some feel that attributing to speakers such further powers of discrimination in relation to such classification questions does not make sense, or that these powers are too unreliable to serve as a basis for theoretically tenable distinctions.

  It is also possible to see such discriminations as not pertaining to the *semantics* of the predications involved, but to certain pragmatic aspects of the use of those predicates.

  Others have argued that at least second order vagueness can be made sense of, but that the distinctions involved are quite different from those relevant to first order vagueness.

  For instance, the question whether a is a second order borderline case of 'bald' might depend on the general dispositions of members of the speech community towards classifying a as bald, or towards classifying a as bald in various contexts.

  When one thinks of second order vagueness as involving criteria that are very different from those for first order vagueness one may be prepared to admit second order vagueness without seeing any reason for accepting third order vagueness (let alone higher orders of vagueness).

7. The Logic(s) of Total and Partial Semantics

- The history of formal logic has for the most part been the history of what constitutes a logically valid inference.
- The fundamental insight on this point was Aristotle's: arguments are logically valid in virtue of their form.
- That also defined the agenda for formal logic:

  (i) articulate the pertinent notion or notions of form and

  (ii) identify which argument schemata (according to the chosen notion of form) qualify as valid.
- Throughout the time from Aristotle to Frege some (but not much) progress was made in the discovery of new notions of form.

  Efforts were made to determine which of the argument schemata generated by the different notions of form are valid schemata but the efforts rarely rose above the level of item-by-item classification, usually with some kind of argumentation why certain schemata were or weren't valid.

- Frege achieved a breakthrough on both fronts:

  (i) he defined a new notion of form (made explicit by the syntax of his *Begriffsschrift*) and (ii) he explicated validity in terms of formal deduction:

  an argument is valid if its conclusion can be obtained from its premises through the (often iterated) application of a handful of formally defined inference principles.
  (Given the immense expressive power of his new notion of form, the number of principles is strikingly small).

- Semantics, in any but an intuitive and informal sense, does not come into this picture.

  (An arguable exception is Venn's account of validity and invalidity of Aristotelian syllogisms in terms of the diagrams named after him.)

  In particular, an explicit semantics for Predicate Logic (the direct descendant of the Begriffsschrift) came only later.

  Crucial in this development:

  (i) The work of Löwenheim and Skolem

  (ii) Tarski's work on truth and logical consequence in the thirties

  (iii) Tarski's development (in the years following World War II) of model theory as a branch of mathematical logic.

  (iv) A hallmark of this development was Henkin's model-theoretic completeness proof for first order predicate logic.

- One of the most important benefits of model-theoretic semantics is that it gives a conceptually clear notion of logical validity:

  We have a strong intuition that a logically valid argument should preserve truth:

  Whenever, and for whatever reason, the premises of the argument are true, then so, of logical necessity, must be the conclusion.

- Within model theory this intuitive notion of validity as truth preserving can be precise in the familiar way:

  for premise set Γ and conclusion B, B is defined to be a *logical consequence* of Γ iff B is true in every model in which the sentences in Γ are true.

- This definition is immensely useful in that it captures the conceptual criterion for validity irrespective of how complex the premises and conclusion may be and irrespective of whether one really ‘understands’ the premises or conclusion or has any direct way of recognizing that premises and conclusion stand in a formal relation that makes the argument valid.

- The value of a completeness proof for first order logic is that it shows us how the semantic notion of logical consequence can be made ‘operational’:

  if an argument is semantically valid, then its conclusion can be obtained from its premises by successive applications of some explicitly specified set of inference rules, which operate according to syntactic principles on syntactic forms;

  if the argument is not valid, then no such derivation is possible.

- The methodology to which this has led can be summarized as follows:

  (a) define a formal language in purely syntactic terms,

  (b) provide a model-theoretic semantics for it that assigns the intuitively right truth conditions to its formulas; and then

  (c) define an algorithmic proof method for that language, with a completeness proof to establish that the algorithm can be used to verify all instances of logical consequence as defined by the semantics, and no others.

  This method has proved immensely fruitful. It has been applied to variants of predicate logic such as Free Logic, to subsystems of the predicate calculus such as Equational Logic, to a wide range of modal logics, including tense logics, multi-agent epistemic logics, dynamic logics and other applications in computer science and so forth.)

- The method has also been applied to the 'logic of vagueness'.

  In fact, we have done a little bit of that here, when 'showing' that classical logic is a viable logic for languages with vague predicates (by defining logical or global consequence as local preservation of truth in supermodels).

  We also implicitly referred to existing completeness proofs when we claimed that the logic of partial models with the truth definition for L given in (1) is Strong Kleene (or Weak Kleene, depending on how we read this definition in the modified setting of partial logic).

- But as soon as we are dealing with partial truth definitions, the semantic method becomes problematic.

One difference is this: For languages with a bivalent model-theoretic semantics there are a number of equivalent ways in which logical consequence can be defined:

(i) as preservation of truth;

(ii) as the absence of counterexamples (models in which all of the premises are true and the conclusion is false), or

(iii) as the converse of truth preservation (whenever the conclusion is false so must be at least one of the premises).

When there is bivalence, these definitions are extensionally equivalent, and this strengthens our conviction that the relation they all define is the intuitively right one.

- When truth definitions are partial, these equivalences no longer hold.

For instance, if we replace the definition of logical consequence given earlier by its converse, the generated logic is no longer Strong Kleene.

This problem gets amplified when we move to ‘multi-point’ structures such as supermodels or Precisification Structures.

We already saw that supermodels suggest at least three different ways of defining logical consequence:

(a) the one that is directly based on the partial truth definition for M yields the Strong Kleene logic;

(b) the global definition, which requires preservation of supertruth, and

(c) the local defintion, which requires local preservation in each of the different complete precisifications.

The last two, we noted give both rise to the same logic, viz. classical logic).

Moreover, when we move from supermodels to Precisification Structures, possibilities multiply further. We will see examples of this later.

- Because of the many different ways in which the model-theoretic semantics for languages with vague predicates can be set up, and because some of these ways allow for a range of different equally plausible ways of defining logical consequence, the landscape of 'vagueness logics' has become increasingly complex.

  There now exist a number of studies that compare these different logics and provide ways of ordering and classifying the different options [a small selection of references: ((Varzi 2007), (Asher, Dever and Pappas 2009), (Cobrero, Egre, Ripley and van Rooij 2012))].

- Because of time limits we have had to abandon a more thorough presentation of these and other logical results about vagueness.

  But we will come back to questions of logic now and then, mostly in connection with the Sorites Paradox.

Literatur

Asher, N., Dever, J. and Pappas, C.: 2009, Supervaluations debugged, *Mind* **118**, 901–933.

Cobrero, P., Egre, P., Ripley, D. and van Rooij, R.: 2012, Tolerant, classical, strict, *Journal of Philosophical Logic* **41**, 347–385.

Fine, K.: 1975, Truth, vagueness and logic, *Synthese* **30**.

Kamp, H.: 1975, Two theories about adjectives, *in* E. Keenan (ed.), *Formal Semantics for Natural Language*, Cambridge University Press, Cambridge, pp. 123–155.

Varzi, A.: 2007, Supervaluation and its logics, *Mind* **116**, 633–676.