1. Begin with a language $L$ containing vague expressions, and assign $T$ to every sentence that is clearly true in $L$ and $F$ to every sentence that is clearly false in $L$. Since $L$ contains vague expressions, some sentences in $L$ will not receive an initial assignment. (Some sentence may fail to receive a valuation for reasons unrelated to vagueness; for example, future contingents or instances of the Liar paradox. We won’t worry about those.)

Now assign either $T$ or $F$ to each sentence that did not receive a valuation on the original assignment. The resulting complete assignment is a precisification of the original assignment. Intuitively, a precisification represents a particular way of making all of the vague terms in $L$ precise. So, for example, one set of precisifications treats 57 grains as the cutoff between heaps and nonheaps; another treats 58 grains as the cutoff; etc. (Precisifications are a species of Fine’s specifications; they are what he calls complete specifications. We will ignore the complexities involved in considering incomplete specifications.)

2. Not all precisifications of a language are admissible. For example, suppose that $A$ is a pile of 43 grains, $B$ is a pile of 44 grains, and $C$ is a pile of 45 grains, and consider the vague sentences:

$A$ is a heap.
$B$ is a heap.
$C$ is a heap.

Here are the values three different precisifications might assign to $A$, $B$, and $C$:


$P_1$ and $P_2$ are admissible (at least, so far as $\alpha, \beta,$ and $\gamma$ are concerned), while $P_3$ is not: no collection of grains goes from a nonheap to a heap by the addition of one grain and back to a nonheap with the addition of yet another grain. It is easy to tell which precisifications are admissible where heaps are concerned. It is not always so easy to say which precisifications are admissible. For example, suppose that Zoe is 13 months old but cannot walk, while Xander is 11 months old but can walk. Is there a constraint on which values an admissible precisification must assign to ‘Zoe is a baby’ and ‘Xander is a baby’? For now, let’s consider this problem solved by supposing that if a precisification is not clearly inadmissible, it is admissible.

3. We are now in a position to define supertruth:

A sentence $\sigma$ in $L$ is supertrue iff every admissible precisification of $L$ assigns $T$ to $\sigma$.

We can define superfalsity analogously:

A sentence $\sigma$ in $L$ is superfalse iff every admissible precisification of $L$ assigns $F$ to $\sigma$.

A sentence that receives $T$ on some admissible precisifications and $F$ on others is neither supertrue nor superfalse; supertruth and superfalsity are mutually exclusive but not exhaustive.
4. Supervaluationism is typically identified with the slogan: “Truth is supertruth.” I have an idiosyncratic concern with this slogan, since if it is correct, we can’t identify the value $T$ assigned to sentences within precisifications with truth. For obviously a precisification can assign $T$ to $\sigma$ without $\sigma$’s being supertrue, so if we identify $T$ with truth and adopt the slogan, a sentence can be true without being true. But if we cannot identify $T$ with truth, we may wonder just what $T$ is. Well (the supervaluationist replies) $T$ is truth on a precisification, while truth simpliciter is supertruth. Yet (I reply) if I can’t help myself to the notion of truth in trying to understand what $T$ is, surely I can’t help myself to the notion of truth on a precisification, can I? Don’t I have to start with truth before getting to truth on a precisification? This issue may be a merely verbal quibble, though I suspect that it is not. I will suppress this concern and adopt the slogan for the purposes of exposition.

5. The supervaluationist says that clearly true sentences in $L$ are supertrue, hence (given the slogan) true, and clearly false sentences in $L$ are superfalse, hence false. Vague sentences in $L$ are neither supertrue nor superfalse, hence neither true nor false. For example:

$$\sigma_1 \quad 1 \text{ grain makes a heap.}$$
$$\sigma_2 \quad 1,000,000 \text{ grains make a heap.}$$
$$\sigma_3 \quad 50 \text{ grains make a heap.}$$

On every admissible precisification, $\sigma_1$ is $F$, so it is superfalse; i.e., false. On every admissible precisification, $\sigma_2$ is $T$, so it is supertrue; i.e., true. And $\sigma_3$ is $T$ on some admissible precisifications and $F$ on others, so it is neither true nor false. As noted above, there may be truth-value gaps arising from sources other than vagueness, so we should not identify vagueness with lack of a determinate truth value (and hence, within the supervaluationist framework, with being neither supertrue nor superfalse).

6. Though supertruth is not bivalent, the supervaluationist can easily respect $\text{LEM}$ and $\text{LNC}$ at the level of supertruth; thus if truth is supertruth, the supervaluationist can respect the logical truth of $\text{LEM}$ and $\text{LNC}$. Assume that the logical connectives have their standard definitions at the level of $T$ and $F$. Then on any precisification, $\sigma \lor \neg \sigma$ and $\neg(\sigma \land \neg \sigma)$ are assigned $T$; thus $\sigma \lor \neg \sigma$ and $\neg(\sigma \land \neg \sigma)$ are both supertrue for any $\sigma$, and hence logically true. There are important questions and controversies about how much of classical logic can be retained at the level of supertruth. Keefe and Williamson both discuss this issue in detail, as will we in the next class.

7. One interesting result of treating truth as supertruth is that the logical connectives are not truth-functional. An $n$-place connective $\chi$ is truth-functional just in case the truth-values of the sentences $[\sigma_1 \ldots \sigma_n]$ determine the truth-value of $\chi(\sigma_1 \ldots \sigma_n)$; in classical logic, the connectives $\neg, \lor, \land, \rightarrow$, and $\leftrightarrow$ are all defined truth-functionally. To see that the supervaluationist connectives are not truth-functional, consider the following two sentences:

$$\alpha \quad (50 \text{ grains makes a heap}) \lor (50 \text{ grains do not make a heap}).$$
$$\beta \quad (50 \text{ grains makes a heap}) \lor (A 13\text{-month-old human is a baby}).$$

Assume that in each disjunction, neither disjunct is true nor false. If we are employing a three-valued logic, each disjunct is assigned the intermediate value $N$. If $\lor$ is truth-functional then since $\alpha$ and $\beta$’s disjuncts have the same values, $\alpha$ and $\beta$ have the same truth-value. But since $\alpha$ is $T$ on all precisifications, it is true, while since $\beta$ is $T$ on some precisifications and $F$ on others, it is neither true nor false. As Williamson points out (146–147), this is a significant advantage for the supervaluationist treatment of vagueness over that given in terms of many-valued logics. For example, consider:

$$\gamma \quad (50 \text{ grains makes a heap}) \lor (50 \text{ grains makes a heap}).$$

According to the three-valued logician, $\alpha$ and $\gamma$ are both assigned $N$, since each of their disjuncts is $N$. But intuitively we only wish to regard $\gamma$ as indeterminate. For the degree-theoretic logician, an even stranger difference results: $\alpha$ has a higher degree of truth than $\gamma$ if and only if the degree of truth of ‘$50$ grains makes a heap’ is less than one-half. The supervaluationist can respect our intuitive attitudes about such cases; apparently neither the three-valued logician nor the degree theorist can.

As we shall see, however, this feature of supervaluationism also gives rise to serious questions about the truth of the supervaluationist’s claim to respect classical logic. Before discussing those questions, however, let us consider how the supervaluationist treats the sorites paradox.

2
8. Consider:

\[ S_0 \] 1 grain does not make a heap.

\[ S_1 \] If 1 grain does not make a heap, then 2 grains do not make a heap.

\[ S_2 \] If 2 grains do not make a heap, then 3 grains do not make a heap.

\[ \vdots \]

\[ C \coloneqq 1,000,000 \text{ grains do not make a heap.} \]

According to the supervaluationist, \( S_0 \) is supertrue, hence true, and \( C \) is superfalse, hence false. But some of the conditional premises are not supertrue, hence not true; thus the argument is unsound. This is because on all admissible precisifications, there is some \( n \) such that \( n \) grains do not make a heap, but \( n + 1 \) grains do make a heap. On each precisification, then, one conditional premise is such that its antecedent is \( \top \) and its consequent is \( \bot \). Thus on each precisification some conditional premise is \( \bot \), so it’s not the case that all of the conditional premises are supertrue; i.e., some of them are not true. Paradox dissolved.

A similar solution is available for the mathematical induction version of the paradox:

\[ S_0 \] 1 grain does not make a heap.

\[ S_n \] For all \( n \), if \( n \) grains do not make a heap, then \( n + 1 \) grains do not make a heap.

\[ C \coloneqq 1,000,000 \text{ grains do not make a heap.} \]

Premise \( S_n \) is a universal generalization. For the reasons just noted, on each precisification one instance of the generalization is assigned \( \bot \). Thus on each precisification the generalization itself is assigned \( \bot \). So \( S_n \) is not only not supertrue, it is superfalse; i.e., false. Thus the argument is unsound. Paradox dissolved.

9. The supervaluationist solution has a troubling consequence. Consider:

\textit{cut} There is some \( n \) such that \( n \) grains do not make a heap and \( n + 1 \) grains make a heap.

\textit{cut} seems to express the claim that there is a sharp cutoff between heaps and nonheaps. Intuitively this claim is false, so we might wish to regard \textit{cut} as false. But the supervaluationist must count \textit{cut} as true for the same reasons she counts \( S_n \) as false. On each precisification, there is some \( n \) such that \( n \) grains do not make a heap and \( n + 1 \) grains make a heap is \( \top \). Thus on each precisification, \textit{cut} is \( \top \), and hence \textit{cut} is true. This is intuitively unpalatable.

Now there is a distinct but related claim that the supervaluationist rejects. Speaking in the metalanguage, we can say that there is \( n \) for which \( n \) grains do not make a heap and \( n + 1 \) grains make a heap true. This is because for each \( n \), at least one precisification assigns \( \bot \) to \( n \) grains do not make a heap and \( n + 1 \) grains make a heap; thus no instance of the statement is true. Thus while it is true that there is a sharp cut-off between heaps and nonheaps (i.e., \textit{cut} is true), there is nonetheless no number of grains for which it is true that that number of grains is the cut-off between heaps and nonheaps. All treatments of the sorities paradox have a counterintuitive consequence; here is the supervaluationist’s. (Some supervaluationists retain the intuitive falsity of \textit{cut}, e.g., Shapiro (\textit{Vagueness in Context}, Oxford 2006), who offers a contextualist version of supervaluationism. But such views are not instances of what we might call orthodox supervaluationism.)

10. One might also wonder about the extent to which the supervaluationist \textit{resolves} the paradoxes. A resolution would not just indicate what makes the argument unsound, but explain why it seems sound. Consider again each premise in the conditional version of the paradox. According to the supervaluationist, many of its premises are not true. But intuitively, each of the premises seems true. What explains this fact? The answer is not obvious.

One possible response the supervaluationist could give is to say that the untrue conditional premises seem true because each is \( \top \) on all but one way of making it precise. Roughly equating the ratio of precisifications on which \( \sigma \) is \( \top \) to those on which it is \( \bot \) with \( \sigma \)'s degree of truth, we could say that \( S_n \) seems true because it has a high degree of truth; it is almost true, which is why we mistakenly regard it as true.
But whatever the merits of this explanation, it cannot be given for the mathematical induction version of the paradox. Employing the rough equation of ratio of precisifications on which $\sigma$ is $T$ to those on which it is $F$ with $\sigma$’s degree of truth, we must say that $S_n$ has no positive degree of truth at all, since it is $F$ on all precisifications. Yet that premise seems no less true than any of its instances. Its apparent truth cannot be explained in terms of its proximity to truth, since it is as far from being true as any sentence can be.

11. One might also wonder about the surprising consequences of supervaluationism that are revealed in its treatment of the sorites paradox on their own terms. $S_n$ is nothing more than the universal generalization of the conditional premises. But $S_n$, unlike all of its instances, is false. It is surprising that a universal generalization should be false when none of its instances is false. It is also surprising that an existential generalization should be rendered true without any true instances, as cut is. In the next class we’ll consider ways a supervaluationist can respond to these surprises, and look at some other troubling consequences of her view.