A Journey into Vagueness

Mike Murphy, CIS 600

May 10, 2006

1 Introduction

Consider an example:

Global Example, Part 1. *Let the "properly rendered" property of this paper be defined as follows:*

1. *If the color of the text in this paper is black, and the background is white, then this paper is "properly rendered" (regardless of the medium: screen or printer).*

2. *If the color of the text is close to black, and the color of the background is nearly white, then this paper is "mostly properly rendered."*

3. *If the color of the text is quite unlike black, or the color of the background is not close to white, then this paper is "not properly rendered."*

Part 1 of the Global Example describes a situation that cannot be suitably represented by classical Boolean logic. Classical logic provides for two truth values: $\{true, false\}$. However, whether or not this paper has been rendered properly has three classes: $\{true, mostly, false\}$. Moreover, the distinction between each of the three classes is not well-defined. The problem contains *borderline cases*, or situations in which there are elements for which it is unclear to what class they belong. In addition to the propriety with which the paper is rendered, the colors "black" and "white" also have borderline cases. Exactly what constitutes the color "white"? Does a pale off-white count?

Propositions or predicates for which borderline cases exist are said to be *vague* [1]. For the purposes of this discussion, a distinction generally is not made between propositions and predicates. Either or both of these logical elements have the capability of being vague. Furthermore, the treatment of vagueness here focuses on properties of objects; thus, the focus will be on predicates that operate on objects. However, for given logical rules and techniques for dealing with vagueness, a proposition may be substituted for a predicate by "wrapping" it in a predicate that returns the truth value of the proposition.

Dealing with vagueness in a logical context is not a straightforward task. There are several different, sometimes incompatible, approaches to working with vague predicates. In this paper, the philosophical background of vagueness, including higher-order vagueness, will be discussed. Following that discussion, three methods of dealing with vague predicates are presented: supervaluationism, fuzzy logic, and a Bayesian approach derived from fuzzy logic. The level of the presentation has been selected to allow the novice, with a basic freshman- or sophomore-level understanding of discrete mathematics, to understand the breadth of the issue of vagueness. For the more advanced reader, the contributions of several principal philosophers and logicians are discussed; works by these researchers may be consulted for a deeper discussion of specific issues. The above Global Example will be carried through the entire paper, in order to provide a deep understanding of *why* the predicates contained in it are vague, as well as to show the differences between approaches to dealing with vagueness. Other examples are provided where appropriate.

2 Philosophical Background

Philosophers deal with three issues that often arise in language: vagueness, ambiguity, and generality. Vagueness specifically deals with the existence of borderline cases in a concept, such as whether or not an object is heavy. Ambiguity refers to the fact that some words may have different meanings, particularly in different contexts. For instance, the sentence "Go by the right path" is ambiguous and requires context to make it into a concrete interpretation. Is the "right" path the "correct" path, perhaps in some kind of moral or ethical context; or is the "right" path the one you take at the junction where you could otherwise take the path that goes to the left? In this sentence, "path" is *ambiguous*: the path may be metaphorical or concrete. If concrete, the path may be quite literally concrete (i. e. made of poured cement and gravel), or the path could be made of dirt or some other material. Thus, the path in this sentence is also *general*, in that the term can describe constructions made of distinct materials, but which have similar properties. [1]

Example 1. *Consider riding in a car with a driver named Dave.*[1] *You are the navigator, selecting the side-roads to use to get to a destination after leaving the interstate. As you approach an important intersection, you warn Dave of an upcoming left turn. At the intersection, you say "turn right here," and Dave quite merrily turns the car... to the right. In trying to get back on course, you find it necessary to make a right turn, and so this time you figure you have it correct: "turn right here." And this time, Dave turns left.*

[1] This is a true story. Despite surviving cancer three times over three decades, Dave unfortunately suffered a fatal heart attack several years after the events described here.

Finally, you learn to give your turning instructions more precisely: "turn left right here" or "turn right, right here." This is an example of another ambiguity present in the word "right" (and a rather extreme reaction thereto). However, the turning directions are not *vague, because there are no borderline cases at the intersections (assuming regular, 2-road intersections at 90 degrees).*

At the crux of the issue of vagueness is the presence of borderline cases, particularly when the exact borders between fallacy and truth (or degrees thereof) are not known or cannot be determined. Consider, for instance, the Sorites Paradox (also known as the Undetermined Boundary Paradox).

Paradox 1. *The Sorites Paradox*

1. *A collection of three or fewer grains of sand does not constitute a "heap" of sand.*
2. *A million grains of sand* is *a heap of sand.*
3. *If $n \in \mathcal{N}$ grains of sand do not make a heap, then $n + 1$ grains of sand also do not make a heap.*
4. *If $n \in \mathcal{N}$ grains of sand DO constitute a heap, then a collection of $n - 1$ grains of sand is also a heap.*

Therefore, if a person starts piling sand on an empty table, one grain at a time, he does not have a heap of sand after adding the first grain. Nor does he have a heap of sand after adding the second or third grains, by rule 1. Furthermore, by rule 3, the addition of a fourth grain does not create a heap, nor a fifth, ..., nor a million. However, rule 2 asserts that a million grains of sand does *constitute a heap. Thus, a contradiction is reached.*

Similarly, if a person starts with a bona fide *heap of one million grains of sand (rule 2) and subtracts one grain at a time, he still has a heap even when no grains of sand remain (rule 4). But rule 1 asserts that three or fewer grains of sand does not make a heap: another contradiction. (Paradox description adapted from [2].)*

The point of the Sorites Paradox is to illustrate that there is obviously some boundary between a heap of sand and something less than a heap. But the division between these two characterizations of a collection of sand grains is unclear and unknown. Thus, a "heap" of sand is a vague concept, which arises in part because of the imprecision of the English language, and partly because a heap of sand is not precisely defined. [1]

Theorem 1. *Any vague predicate can be used to instantiate a paradox of contradiction that is reducible to the Sorites Paradox.*

Proof. Consider any vague predicate q. By definition, q is vague, so it must have a borderline case. Furthermore, the borderline case of q is assumed not to be known, since q could be split into separate sharp predicates in such a situation.

1. Instantiate a base case of the sub-predicates of q such that the instance is decidedly false. Call this instance i_0.

2. Specify the sub-predicates for some instance of q such that the instance is unquestionably true, and call this instance i_ω.

3. Assert that, if some setting of the sub-predicates of an instance of q is such that the instance is false, then one "increment" of the sub-predicates does not change the truth value (i. e. assume the border is not crossed by "incrementing" the sub-predicates). Note here that "increment" is quoted to avoid implying that the operation is always an addition of 1, as it is in the Sorites Paradox.

4. Start from the base case in step 1 and gradually "increment" the underlying cases that change the sub-predicates, until the sub-predicates match those of step 2.

5. At the base case, i_0 is false (trivially, since it was defined that way).

6. For some i_k reached by "incrementing" from i_0, assume that i_k is false. By the rule given in step 3, i_{k+1} is also false.

7. When the iteration reaches i_ω, i_ω must be false by the previous steps. But i_ω is true by step 2: contradiction.

Substitute the heap problem from Paradox 1 for q, and the result is exactly the Sorites Paradox. (Theorem and proof inspired by Sorensen's proof from [3].) □

We can make use of this theorem to show that the predicates in the Global Example are vague.

Global Example, Part 2. *Consider the rules given in part 1 of the Global Example. One way of expressing the color of the text or the background is to use a 24-bit hexadecimal RGB (Red-Green-Blue) value that gives the intensity of each component color. Pure white is the maximum intensity of all three component colors (*0xffffff*), while pure black is the zero of all intensities (*0x000000*).*

Rewrite rule 1 to say that the paper is properly rendered if the text has the color represented by 0x000000 *and the background is represented by* 0xffffff. *Start with these two color representations and add one level of intensity to each component color of the text, and reduce by one level of intensity each component color of the background (in* *other words, repeatedly add* 0x010101 *to the text color, and repeatedly subtract* 0x010101 *from the background color).*

After one step in the iteration, the text has the color represented by 0x010101*, and the background may be represented by* 0xfefefe*. These colors are probably indistinguishable from pure black and pure white on most output devices, so consider that rule 2 is satisfied. Moreover, assume that if rule 2 is satisfied on this iteration, one more small adjustment to the color is not going to remove the "mostly properly rendered" conclusion. But once the iteration finishes and the colors are inverted (text is representable by* 0xffffff *and* 0x000000 *represents the background), rule 3 is clearly satisfied, contradicting rule 2. This is an instance of a contradiction reducible to the Sorites Paradox, and so the propriety predicate is vague.*

In 1985, Sorensen extended a version of the Sorites Paradox to show that the concept of vagueness is itself vague. His argument, presented below as a theorem, has been met with some resistance from philosophers such as Hyde, Tye, and Deas. A discussion of the controversy, and the philosophical resolutions to it may be found in [3], from which this theorem and proof were taken.

Theorem 2. *(Higher-Order Vagueness): The concept of "vagueness" is vague.*

Proof.

1. Start with an instance of the Sorites Paradox on the "smallness" of a natural number, creating the vague predicate *small*:
   (a) Assert that 1 is small: $small(1)$
   (b) Assert that 10^{10} is not small: $\neg small(10^{10})$
   (c) $\forall n \in \mathcal{N} : small(n) \implies small(n+1)$
   (d) By induction starting from 1, $small(10^{10})$ is true: contradiction.
2. Now define a predicate *n-small*:
   (a) $\forall k, n \in \mathcal{N}$ *n-small(k)* $\iff small(k) \vee (k < n)$
   (b) *1-small* reduces to *small* and is therefore vague; however, for some large (i. e. non-small) value of n, such as 10^{10}, the predicate *n-small* is not vague because the predicate reduces to a less-than operation since $small(n) = false$.
3. The proof follows from another instance of the Sorites Paradox:
   (a) *1-small* is a vague predicate
   (b) 10^{10}-*small* is not a vague predicate
   (c) $\forall n \in \mathcal{N}$ if *n-small* is vague, then *(n+1)-small* is also vague.

(d) By induction starting from 1, it follows that 10^{10}-*small* is vague: contradiction.

(Proof adapted from [3].) □

Since the concept of vagueness is itself vague, the methods of dealing with vagueness will not be satisfactorily clear. Moreover, the method that produces the "most desirable" result (another vague notion) will depend on the problem. Three such methods will now be discussed, starting with supervaluationism.

3 Supervaluationistic Approach

One method of dealing with vagueness in logical situations is a philosophical technique called *supervaluationism.* In this paradigm, truth values do not apply to borderline cases. Rather, the supervaluationist tries to "precisify" a vague predicate. This notion follows from the concept of "hyper-ambiguity" proposed by Fine and Lewis, which posits that a vague concept is in reality an extremely large collection of similar precise concepts. Thus, a true vague statement is a statement that comes out true under any disambiguation of its meaning. Moreover, disambiguation is required before a truth value may be assigned to a vague predicate. [1]

With supervaluationism, certain logical tautologies, such as $p \vee \neg p$, hold in the same manner as they do in classical logic. For instance, supervaluationism allows one to say, "either this paper is properly rendered or it is not properly rendered" and thereby express a tautology. However, the exclusion of truth values for borderline cases means that a proof that a statement is not true is not sufficient to say that the statement is false. [1]

Another consequence of the supervaluationist approach is that operations on supervaluated predicates match classicial logic operations only when each vague term in the sentence is interpreted with the same meaning. Linguistic ambiguities might in fact *require* different interpretations of the same word in the same sentence. Consider expressions like "She's having a ball at the ball" or "non-toxic ant poison" [1]. The term "ball" has two distinct interpretations in the same sentence: a "good time" and a "dance," respectively. Similarly, "non-toxic ant poison" that is not toxic to the ant will not be a good investment; in this case, the concept of toxicity is implicitly dependent on the affected species. In these cases, the logical operators in supervaluationistic logic do not match those of classical logic. [1]

Global Example, Part 3. *The precisifications of the vague predicates in part 1 of the Global Example depend on the representation of the colors and on the output device on which this paper is rendered. Consider the 24-bit representation of color described in* *part 2 of this Global Example. Most modern computers are capable of representing 24-bit color; however, back in the Bronze Age (c. 1995), some computer systems could only represent 256 or fewer colors. Representations using 256 colors used three bits each for red and green, but only two for blue, owing to color sensitivity differences in the human eye [4].*

Possible degree curves *(more on these in the next section) that relate the "level" of truth to the vague predicates and to the possible output devices are shown in figures 1 to 4 (found in the next section). The x-axes of these graphs represent a linear transformation of the hex color values to a scale that runs from totally black to totally white. Thus, moving to the right in each graph is moving toward "more white" values, while moving to the left is moving toward "more black" values. Scale labels have been omitted from the x-axis to avoid confusion.*

Consider a text color that is at the first tick mark along the x-axis, moving right from the origin. If the color of the text at that point is precisified for a rendering device using 8-bit color (figure 1), then the text is still within the range of the "mostly properly rendered" rule. However, if the color value is precisified for a device that supports 24-bit color, then the paper will not be properly rendered, because the 24-bit black color degree curve (figure 2) shows a truth probability of zero. A similar situation occurs with the white background (figures 3 and 4).

Now jump in a time machine and go back to the Stone Age (say, 1975), when printers were either glorified typewriters or daisy-wheel devices that could only print in black. Assume that printer paper was only available in white. Under this universe of discourse, there is only one possible precisification for each color in the paper. Thus, rule 1 is always satisfied. In this situation, the predicates "the text is black" and "the background is white" are said to be supertrue *because they are true under all possible precisifications of the vague concepts in this particular universe. [1]*

4 Fuzzy Logic Approach

Another method of handling the issue of vagueness is to treat vague predicates as having continuous ranges of values, in order to make the representation of such predicates more "human-like." Using a continuous range of possible truth values allows predicates to have degrees of truth and fallacy, instead of forcing predicates to be completely true or completely false. In other words, fuzzy logic maps logical truth and fallacy onto the real numbers in the range $[0, 1]$, instead of onto the standard Boolean set $\{true, false\}$. [5]

Adding extra logical values to the two-element Boolean set was first proposed by Jan Lukasiewicz in 1920, when he added a third "unknown" truth value. Other prominent mathematicians and computer scientists in the early 20th century also researched multivalued logic, including Emil Post, Stephen Kleene, and Donald Knuth. In 1965, Lotfi Zadeh introduced the concept of fuzzy logic, with an infinite number of possible logical values. [5]

Reasoning with fuzzy logic differs considerably from classical Boolean reasoning, largely because fundamental Boolean principles such as the Law of Excluded Middle ($q \vee \neg q = true$) break down [6]. Fuzzy logicians view classical reasoning as a subset of fuzzy reasoning, which occurs only at the limits $((p(q) = 0) \vee (p(q) = 1))$. Furthermore, predicates from any logical system can be converted to real degrees in $[0, 1]$, or "fuzzified." [5]

The choice of the range $[0, 1]$ is significant in fuzzy logic because it allows for the use of statistical methods in fuzzy reasoning. Consider a probability evaluation function p:

$$p : P \longrightarrow \mathcal{R}_{[0,1]} \tag{1}$$

That is, the function p takes an element from the set of predicates P in the universe of discourse and returns its *degree* (or probability) of truth, in the real interval $[0, 1]$. Notice, however, that an ambiguity arises here: does $p(x)$ mean the probability that x is true (in the classical sense), or does $p(x)$ mean that x is $[p(x)]\%$ true? The application of fuzzy logic does not make a clear distinction between these two cases.

From definition 1, the logical negation of some predicate q can be defined according to probabilistic negation: $p(\neg q) = 1 - p(q)$. If two predicates are conjoined, the degree to which the resulting conjunction is satisfied is defined to be the minimum of the probabilities of the two conjuncts. For disjunction, the maximum of the probabilities of the disjoined terms is taken. Equations 2 and 3 summarize these rules. [7]

$$p(q \wedge r) = min(p(q), p(r)) \tag{2}$$

$$p(q \vee r) = max(p(q), p(r)) \tag{3}$$

There is an important distinction between the input propositions (or predicates) q and r, and the output of a fuzzy logic equation such as equation 2 or 3. The values $p(q)$ and $p(r)$ represent *degrees of belief* in the truth or fallacy of q and r. On the other hand, the value of something like $p(q \wedge r)$ represents a *degree of satisfaction* (in this case, the degree to which the conjunction has been satisfied). These two measures are conceptually different, although fuzzy logical operations tend to blur this distinction. [7]

Only the rules for negation, AND, and OR are defined according to probabilities. Implication is defined using a method that more closely resembles imperative operations. Several such constructions have been proposed, but Brachman and Levesque (whose presentation this paper is following) chose Mamdani's procedure, which is given below: [7, p. 256-258]

1. Use the degree curves for each predicate to transform the inputs from vague statements to probabilities. That is, compute $p(z)$ for all $z \in P$.

2. Evaluate the antecedents by combining the probabilities according to the rules given in formulas 2 and 3.

3. Evaluate the consequents by selecting outputs from the degree curves for each consequent. There are multiple ways of making this selection.

4. Aggregate the consequents by combining the possible values in such a way as to take into consideration the degree of each antecedent. Again, there are multiple ways to do this step.

5. "Defuzzify" the output by using the aggregated degree curve to calculate a weighted average value for the output. The weighted average can be computed in one of several different ways, such as taking the center of area of the curve.

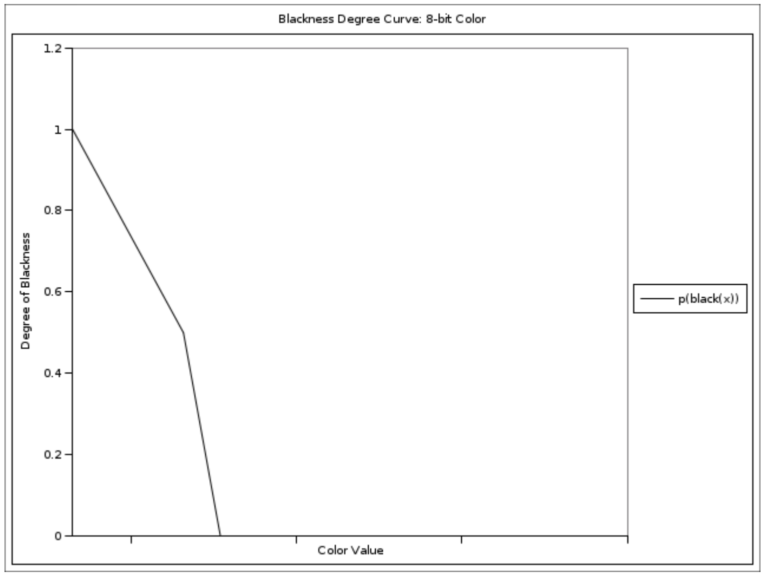
The description of Mamdani's approach is rather vague and allows for a number of decisions to be made about the specific operations to be performed at each step. This situation arises because there are several different variants of this particular fuzzy logic approach, and the implementer of fuzzy logic operations may choose which variant to use in order to match the specific problem. However, the implementer should be cautioned that different variants may be incompatible with each other. [7] Here is one possible fuzzy logic solution to the Global Example problem:

Global Example, Part 4. *One possible fuzzy logic method for dealing with the vagueness in the rules of the Global Example is to use the 8-bit degree curves (figures 1 and 3) in step 1 of Mamdani's method. Assume that we are told the color of the text is equivalent to the first tick from the left on the x-axis in figure 1, and the background color is equivalent to the first tick from the right, using figure 3. We apply Mamdani's method as follows:*

1. *Traveling up from the x-axis at these two points, the values of the curves are approximately 0.7 and 0.75, respectively.*

2. *Rules 1 and 2 of the Global Example (see part 1) rely on the conjunction of the two color probabilities. Thus, we take the minimum according to equation 2 and get 0.7.*

Figure 1: Degree Curve for Blackness (8-bit Color)



3. *One way to evaluate the consequent is to draw a horizontal line across its degree graph (figure 5) at the calculated value of the combined antecedents: 0.7. The curve intersects such a line at about the point on the x-axis labeled with category number 5.*

4. *We can skip the aggregation of consequents because there is only one consequent.*

5. *Let category 6 in figure 5 represent satisfaction of rule 1 in the Global Example, and let categories 2 through 5 represent satisfaction of rule 2. We conclude that the paper is "mostly properly rendered" in this case.*

It is important to realize that the result of applying fuzzy logic by a rule such as Mamdani will give different answers for different choices made at each step. For instance, had we selected the 24-bit degree curves, the antecedent values would have both been zero, and the conclusion would have been that the paper was not properly rendered.

5 Bayesian Approach

One approach to solving a problem involving vague predicates is to cast the problem into a solution space that does not have borderline cases. Degrees of belief and satisfaction can be cast as *subjective probabilities* by means of Bayes' Rule. While some

Figure 2: Degree Curve for Blackness (24-bit Color)

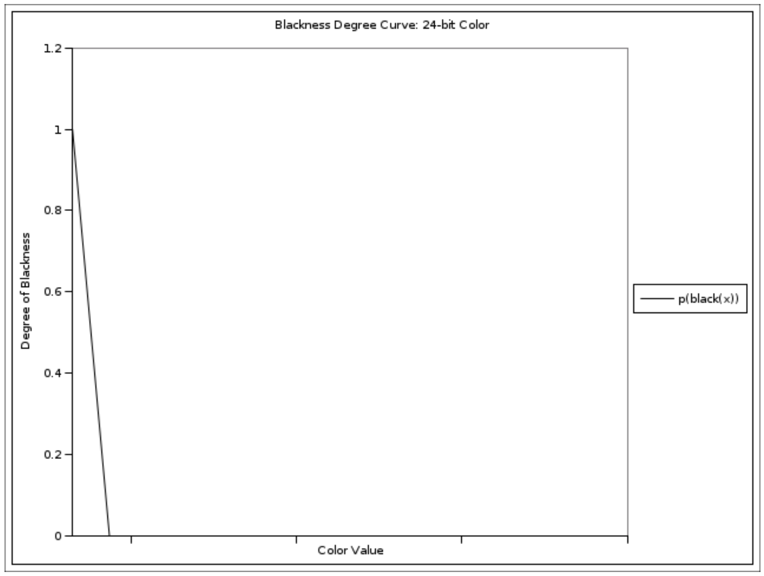


Figure 3: Degree Curve for Whiteness (8-bit Color)

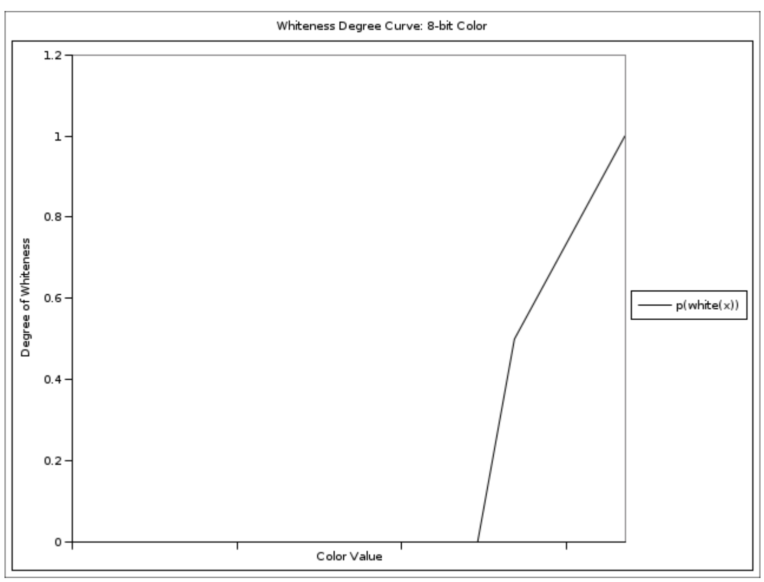


Figure 4: Degree Curve for Whiteness (24-bit Color)

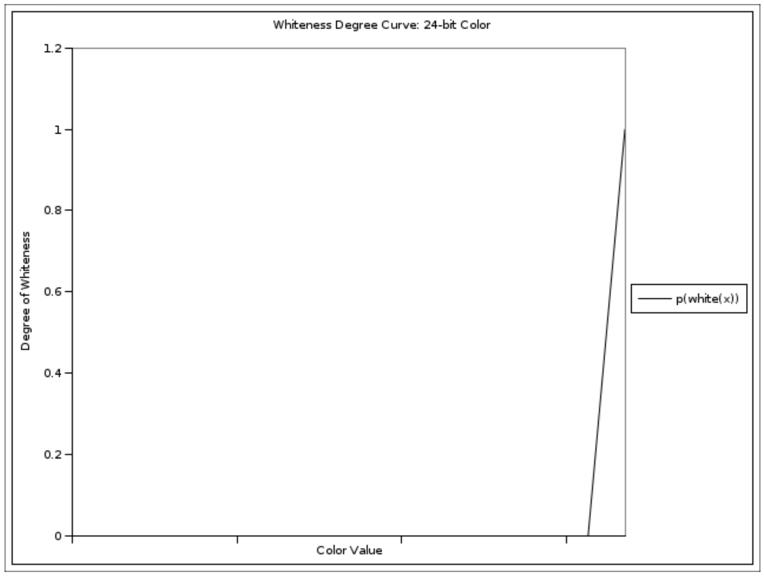
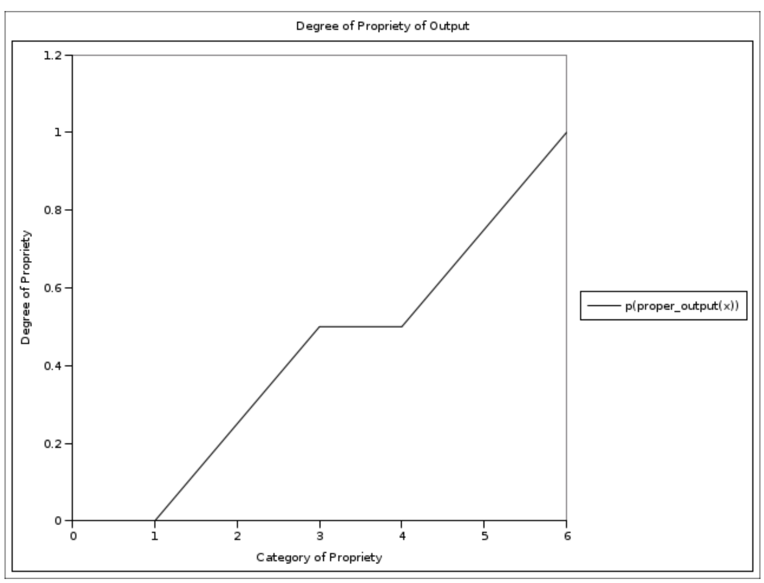


Figure 5: Degree Curve for Propriety of Rendering



aspects of this procedure are still dependent upon the specific problem to be solved, the Bayesian approach avoids some of the incompatibilities in the different fuzzy logic approaches to resolving vague sentences. [7]

If the probabilities of two events a and b are not independent of each other, then the following formula (Bayes' Rule) relates the probability of event a given b to the probability of event b given a: [7, p. 242]

$$p(a \mid b) = \frac{p(a) \times p(b \mid a)}{p(b)} \tag{4}$$

Example 2. *One good example of the way in which Bayes' Rule works comes from Brachman and Levesque: let "a" represent a disease while b represents a symptom. If a person exhibits the symptom b, the probability that he or she has the disease ($p(a \mid b)$) is given by equation 4. In order to calculate this probability, we must know the probability of diagnosing the disease in the population ($p(a)$), the probability of seeing symptom b in the population regardless of the actual cause ($p(b)$), and the probability that disease a will produce symptom b ($p(b \mid a)$). [7]*

To use Bayes' rule to resolve a vague sentence, a function g is needed in place of the predicate a. The area under g should be 1, and g is to be defined in such a way that $g(i) = p(m(i) = n)$ for all values of n in the domain of a *base measure* function m. The base measure function $m(i)$ is associated with the predicate a applied to the individual i and essentially functions as a descriptor for some property. [7]

As was the case in fuzzy logic, the Bayesian approach (with non-independent predicates q and r, and some condition y) requires special definitions for conjunction and disjunction: [7]

$$p(q \wedge r \mid y) = min(p(q \mid y), p(r \mid y)) \tag{5}$$

$$p(q \vee r \mid y) = max(p(q \mid y), p(r \mid y)) \tag{6}$$

The primary limitation of the Bayesian approach to solving actual problems is that *a priori* information about $p(a)$ and $p(b)$ is needed. In example 2, one would need information about the prevalence of both the disease and the symptom in the population. To solve vague predicate problems using this method, certain assumptions are often needed because these probabilities are not known. [7] Once again, the assumptions taken affect the final results, and so the Bayesian approach is not a panacea for solving a vague problem.

Global Example, Part 5. *Let a be the predicate that the paper is* at least *"mostly properly rendered", and let b be the conjunction of the predicates that the text is black and the paper is white. We happen to know from part 4 and the degree curves that the probability that the text is black is 0.7, the probability that the paper is white is 0.75, and the conjunction of these probabilities is 0.7 (assuming 8-bit color).*

But the conjunction of these probabilities is *the independent probability that the paper will be at least "mostly properly rendered," since the predicate a is not independent of the predicate b. Furthermore, if we are given the predicate a to be true, then $p(b \mid a) = 1$. So applying the Bayesian formula, we get:*

$$p(a \mid b) = \frac{p(a) \times p(b \mid a)}{p(b)} = \frac{0.7 \times 1}{0.7} = 1$$

The end result is that this paper is almost surely at least "mostly properly rendered," given an 8-bit color environment and a priori *information about the actual encodings of the text and background colors. Notice that this result does not give any more information than the fuzzy logic approach gave for this problem. Also notice that without the* a priori *information we had, the Bayesian approach would be essentially useless for this problem.*

6 Conclusion

In this paper, philosophical issues related to vagueness were discussed, and several approaches to dealing with vagueness were considered. Vague predicates are those that have borderline cases of truth, and the concept of vagueness itself is vague because it has borderline cases. As a result, the methods for dealing with vagueness are dependent upon assumptions and specific decisions that must be made by the implementer.

There is no escape from vagueness in the world. Rarely do issues or measurements fall neatly into one of two Boolean truth conditions. By understanding the philosophy behind vagueness, along with a few of the popular computational methods used to deal with it, it is possible to design systems that can operate in a world full of borderline cases. Sensor networks, agent-based systems, robots, and artificial intelligence systems are just a few of the types of computer applications that must be able to function in a vague world.

References

[1] Roy Sorensen. Vagueness. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Stanford University, 2003. URL `http://plato.stanford.edu/archives/fall2003/entries/vagueness/`.

[2] (community edited). Paradox of the heap. In *Wikipedia*. Wikimedia Foundation, Inc., 2006. URL `http://en.wikipedia.org/wiki/Sorites_Paradox`.

[3] Achille C. Varzi. Higher-order vagueness and the vagueness of 'vague'. *Mind*, 112:295–298, 2003.

[4] (community edited). 8-bit color. In *Wikipedia*. Wikimedia Foundation, Inc., 2006. URL `http://en.wikipedia.org/wiki/8-bit_color`.

[5] Pragya Agarwal. Lotfi zadeh: Fuzzy logic-incoporating real-world vagueness. In *CSISS Classics*. University of California Santa Barbara, 2006. URL `http://www.csiss.org/classics/content/68`.

[6] Eric W. Weisstein. Law of the excluded middle. In *MathWorld*. Wolfram Research, Inc., 2006. URL `http://mathworld.wolfram.com/LawoftheExcludedMiddle.html`.

[7] Ronald J. Brachman and Hector J. Levesque. *Knowledge Representation and Reasoning*. Elsevier, Inc., 2004.