By speaking of the bald man, I am of course referring to the most clear-cut of the paradoxes of vagueness, the sorites paradox. Or, strictly speaking, I am referring to one of the dramatizations of this paradox. This case is nevertheless fully representative of the general issues involved. (For the sorites paradox in general, see e.g. Keefe and Smith 1987 or Sainsbury 1995, ch.2.) The allegedly paradoxical argument is well known. It might be formulated as follows:

(P.1) Premise one: a man with no hairs is bald.
(P.2) Premise two: If a man with n hairs is bald, then so is a man with n + 1 hairs.
(C) Conclusion: No matter how many hairs a man has, he is bald.

The paradox lies in the fact that the premises (P.1) and (P.2) seem to be unproblematically true but the conclusion (C) false. Hence it is in order to have a closer look at the logical form of the argument that gives rise to the conclusion.

Now the logical form of the statements involved here seems to be the following (using the conventional first-order notation):

(P.1)* \((\forall x)(H(x,0) \supset B(x))\)
(P.2)* \((\forall n)(\forall x)(H(x,n) \supset B(x)) \supset (\forall x)(H(x, n+1) \supset B(x))\)

(C)* \((\forall n)(\forall x)(H(x,n) \supset B(x))\)

Theses in turn exemplify the members of the schema of mathematical induction.

(P.1)** \(A(0)\)

(P.2)** \((\forall n)(A(n)) \supset A(n+1))\)

(C)** \((\forall n)A(n)\)

The inferences from (P.1)* – (P.2)* to (C)* and from (P.1)** – (P.2)** to (C)** are valid in the usual first-order logic plus the principle of mathematical induction. Since neither of these inferences seems to be subject to doubt, we seem to have a serious paradox facing us. As was pointed out, the premises (P.1) and (P.2) seem true while the conclusion (C) is a (bald?) falsehood. Yet the inference turns out to be valid when it is construed in an apparently completely unproblematic way. As was pointed out, the inferences to (C)* and to (C)** from their respective premises are instances of mathematical induction. This mode of reasoning is thus involved in the problems of vagueness, and has in fact been discussed in connection of these problems. But it seems to be beyond any reasonable doubt, and in fact no serious doubts seem to have been leveled at it.

If the reasoning involved in the paradox seems so unproblematic, how can the paradox be explained away? What can the bald man tell us? Here comes the first and quite possibly foremost strategic moral that can be elicited from discussions of the different paradoxes such as the sorites paradox and the paradox of the liar. An enormous amount of inspiration and perspiration has been devoted to sundry “solutions” of the
paradoxes. It is nevertheless by this time eminently clear that the only way of clearing up such major paradoxes is a deeper analysis of the basic concepts figuring in them. To discuss the paradoxes as separate puzzles without digging deeper into their sources in logic and semantics is not much more instructive than to solve crossword puzzles. The difficulty of the paradoxes is reflected in the fact that this proposed deeper analyses forces us to take a long hard look at the questions that underlie all logic. In the case of the sorites paradox the basic concepts that hold the key to understanding the paradox are negation and mathematical induction. (For negation, see Hintikka 2002.)

One way of coming to see this is to look at the attempts that have been made to “solve” the paradox. What are they like? One’s first intuitive reaction to the puzzle of the bald man is undoubtedly to say: But baldness is not a black-or-white concept. In some cases — for some numbers of hairs — it does not make sense to say that a man is bald or that he is not. Technically, this is expressed by saying that the predicate “bald” exhibits truth-value gaps. Surely that is why we do not really have to worry about the paradox.

This line of thought has given rise to attempts to deal with the paradox by postulating truth-value gaps for predicates like “bald”. In other words, it is assumed that our basic nonlogical predicates are not defined for some argument values, e.g. that $P(x)$ is not defined for the value $b$. Then $P(b)$ is neither true nor false. In the case of the paradox dealt with here, the vague concept will obviously be the notion of baldness.

Unfortunately but perhaps not unexpectedly, this line of thought does not yield a genuine solution to the paradox, at least not without a great deal of further ado. This further work has among other things led to the theory of supervaluations. Even without
discussing the details of such theories, it is safe to say that they have not led to a definitive solution of the sorites paradox in spite of the plausibility of the idea of truth-value gaps. Without going into any details, a sample difficulty will hopefully give you a sense of the difficulties here. Suppose that we allow the predicate "bald" to have truth-value gaps. But let us then consider the predicate $B^*(n)$ that says that a man with $n$ hairs is either bald or neither-bald-nor-not-bald. Then an obvious variant of the inductive argument above apparently proves that no matter how many hairs a man has, he is either bald or neither-bald-nor-not-bald, which is as absurd a conclusion as the original one.

There does not seem to be any way out of this predicament. The postulation of truth-value gaps was intended to eliminate any sharp jump from build to not build by addition of one hair. But the very same postulation creates a similar sharp jump from cases of baldness to cases of neither-bald-nor-not-bald, thus re-creating the paradox.

Thus we have here a second paradox in our hands. It lies in the fact that the eminently plausibly truth-value gaps idea does not immediately solve the sorites paradox. Now the reason for this second paradox was already diagnosed above. It lies in reality in the superficial attitude to what the logic is that is involved in the sorites paradox. We must dig into the most basic questions here. What is the job description of logical words? It is thought far too often that their task is to serve as guideposts to inference. If so, you could presumably change their meaning by changing the inference rules that govern them.

This view is fundamentally misguided. What the real job of logical words is, is to serve the representation of reality in language. All communication and all inference relies on such representation. It is in virtue of their representational role that logical
words can serve to mediate inferences. And this representational role is not played by means of formal rules of inference. Hence simply to postulate changes in the formal rules of inference or in truth-tables does not automatically result in interpretable language. Likewise changes in the customary truth tables do not automatically have a semantical interpretation. Postulating an indefinite truth value might have some plausibility in the case of a notion like baldness which admits of differences in degree. But in other cases an attribution of an indefinite truth value has no meaning whatsoever. I know what it is for an object to be yellow and likewise what it is for it not to be yellow. But if you tell me that your shirt is neither yellow nor not-yellow, I can associate a meaning to your utterance only by assuming that you are using language in a nonliteral sense. Likewise there is no limbo between existence and non-existence.

The only way out seems to change our concept of negation from the contradictory negation to something else. This means for instance in one of our examples assuming that there are separate criteria of deciding whether it is yellow and of deciding whether it is not-yellow. Naturally, this latter option would mean using negation in some way different from the familiar contradictory negation.

The bald man paradox is not a good example here because in it we are tempted (or at least willing) to think of baldness and hirsuteness as separate concepts. Then the truth-value gaps would be simply the cases intermediate between these two independently identifiable concepts. But if they are separate concepts, the use of negation in (P.1)*-(P.2)* and (C)* is not correct.

More generally, one cannot simply postulate truth-value gaps or postulate unconventional truth-tables and expect to obtain an interpretable language. Language is
holistic in that different logical constants do not do their representational job in isolation but always in cooperation with each other and with other notions. Physicists have apparently had more semantical common sense than philosophers here. They have been seriously puzzled by problems like the Schrödinger’s cat but they have not followed Reichenbach’s advise and tried simply to stuff the poor beast into a truth-value gap.

I agree that the postulation of truth-value gaps is a step in the right direction. However, it does not suffice alone to clarify the problem situation. Rather, the introduction of truth-value gaps is like the discovery of a brilliant new gambit in chess that nevertheless does not work unless the overall rules of the game were changed.

But what is the counterpart in such a change of rules in the case of the game of logic? The postulation of truth-value gaps means that atomic sentences with suitable nonlogical predicates are neither true nor false, in other words, that the law of excluded middle holds for them. But you cannot change your logic only in its application to some particular subject matter. Truth-value gaps for atomic sentences force you to acknowledge truth-value gaps for other propositions, including quantified ones, thus making the interpretation of the language even more difficult. What would the reality have to be like in order to allow us to make sense of a world where mammals neither exist nor do not exist?

Hence the philosophers who have postulated truth-value gaps ought not have used the usual textbook logic. If we change our logic as it applies to atomic propositions, we must change it across the board. Thus one of the things that the bald man tells us is that it simply is a mistake to give up the law of excluded middle for some predicates but otherwise use the conventional first-order logic which presupposes the law. But logicians
defending truth-value gaps have never explained what an alternative logic might be and how it is to be interpreted. Yet only such an interpretation could make clear sense of even our familiar logical truths. For instance, if vague predicates are allowed, the following sentences will not any longer remain logically true even though they must be taken to be logical truths according to our usual logic employed in the usual attempts to solve the sorites paradox:

(1) There is someone such that if he or she can solve this problem, anyone can

(2) (Peirce’s paradox.) Someone will commit suicide if he fails in business if and only if someone will commit suicide if everyone fails in business (See here Hintikka 2006.)

These apparently have the respective logical forms

(1)* (\exists \, x) (A(x) \supset (\forall \, y) (A(y))

(2)* (\exists \, x) (F(x) \supset S(x)) \leftrightarrow (\exists \, x) ((\forall \, y) F(y) \supset S(x))

Yet (1)*-(2)* cannot be correct representations of the logical forms of (1)-(2), for (1)*-(2)* are logical truths of the usual first-order logic while (1)-(2) are obviously not. Hence it is inappropriate in dealing with the truth-value gaps idea to employ the received first-order logic. As (1)-(2) illustrate, it does not even yield right results in all cases.

Another reason is that when logic is formalized, it becomes very awkward to maintain a distinction between laws that hold for atomic formulas and others. For atomic formulas are typically used in formalizations as placeholders for arbitrary formulas. Yet in the
logic used in the usual discussions of paradoxes no explanation is given of how truth-value gaps of complex predicates involving quantifiers could be interpreted.

The basic insight that is not heeded in the earlier discussion is thus that if we admit truth-value gaps we must have in our logic a strong dual negation that does not obey the law of excluded middle. This is perhaps the most important thing that the bald man can tell us. Otherwise, we for instance cannot deal with the totality of truth-value gaps nor with the meaning of any third (indefinite)-truth-value.

Fortunately, there exists a logic which can serve these purposes. This logic is known as independence-friendly (IF) logic. (For it, see e.g. Hintikka 1996 and 2004.) It is more fundamental than our ordinary first-order logic, which could be called, somewhat inaccurately historically, the Frege-Russell logic. As has been repeatedly pointed out, Frege-Russell logic is unnecessarily restricted in its expressive power. There are several crucial logical and mathematical ideas that cannot be expressed in the ordinary first-order logic but can be expressed by means of IF logic, including equicardinality, infinity and topological continuity. Now we can add to the services IF logic can perform for the solution of the sorites paradox.

The source of this additional representational power is purely logical. Part of the semantical (representative) function of quantifiers is to express the dependence and independence of variables. A bound variable $y$ depends (in the interpreted real-life sense) on another variable $x$ if and only if it is bound to a quantifier ($Q_y$) that depends formally on the quantifier ($Q_x$) to which $x$ is bound. Now in Frege-Russell languages such formal dependence is expressed by the nesting of scopes. But the scope relation is of a very special kind, among other features transitive and antisymmetric. Hence it cannot serve to
express patterns of dependence relations that do not all have all these features. Hence, it cannot do its whole job.

This restriction is removed in IF logic which thus can be formulated notationally by allowing a quantifier (Q₂y) to be independent of another one, say (Q₁x), even though it occurs in its (formal) scope. This is indicated by writing it (Q₂y/Q₁x) (For IF logic, see e.g. Hintikka 1996, 1998 and 2002(b).) A semantics is thus obtained for IF logic from the usual game-theoretical semantics for ordinary first-order logic by letting every move prompted by an independent quantifier (Q₂y/Q₁x) to be informationally independent of the move prompted by (Q₁x). The truth of a sentence S is defined in the usual way as the existence of a winning strategy for the verifier in the corresponding semantical game G(S). The falsity of S is defined as the existence of a winning strategy for the falsifier.

It turns out that even when all the rules for semantical games are otherwise the usual (“classical”) ones, some games are not determinate: neither the verifier nor the falsifier has a winning strategy in G(S). The S is neither true nor false. The law of excluded middle does not hold in IF logic.

This conclusion is perfectly straightforward. There are nevertheless several remarkable things about it. First, IF logic ensues inevitably from purely classical semantics (classical rules for semantical games). The only novelty is the admission of independence. This feature is needed for IF logic to have its required expressive power. IF logic is thus not a “nonclassical” logic. It should rather be called neoclassical or perhaps hyperclassical logic.
A further point is especially important here. Not only does IF logic lead us to acknowledge sentences that are neither true nor false. Let’s all them indefinite. The game-theoretical semantics for IF logic shows what it means for a sentence S to be indefinite. The truth of S in a model (“world”) M means the existence of a naming strategy for the verifier. This is a definite combinational (set-theoretical) fact about M. The same holds for the falsity of S. By the same token the fact that S is indefinite means that M has certain objective combinatory features. If I know that S is indefinite, I know that the world has those features. Hence game-theoretical semantics assigns an objective meaning to all attributions of a truth-value, including the indefinite one.

At this point, a look at the wider context of the discussion here is in order. The possibility of making interpretational sense of attributions of indefinite truth-value does not only clarify the meaning of formal logics without the law of excluded middle. It opens new avenues of the applications of logical methods. For instance, why is the axiomatic method so useful? In it we study a class of all possible scenarios which are models of the given system. Instead of having to go out to the world and investigate such possible scenarios experimentally and observationally, an axiomatist can study them by drawing logical consequences from the axioms, without needing more than pencil and paper (and, in these days, a computer). This advantage is even independent of whether the logical consequence relations needed are themselves mechanizable.

Now usually the class of models (“scenarios”) so examined are the ones in which the axioms are true. But we could equally well study different classes of models in the same way, for instance the class of models in which the axioms are not false. In mathematical logic, this strategy has in fact been employed under the title “no-
counterexample interpretation”. This interpretation has been used in an interesting way by Georg Kreisel: see Feferman (1996). Now we can see its significance in a wider context.

The reader might have wondered about what was said about (1)*-(2)* above. If the only difference between received first-order logic and IF first-order logic lies in allowing independence by means of the slash notation, how can the switch from the former affect the status of formulas like (1)*-(2)*? The do not contain any slashes. The answer is that we have to use IF logic not only when slashes are present but also as soon as the basic nonlogical predicates allow for truth-value gaps, that is, violations of the law of excluded middle. This puts in perspective what has been said earlier in this paper. We can now see the precise technical source of the difficulties into which the theorists relying on truth-value gaps have run into. Their mistake is not to have realized that the use of IF logic is necessary not only when slashes are present but also when the primitive predicates are assumed to have truth-value gaps. They have continued blithely to use the traditional first-order logic.

The logical laws of IF first-order logic differ essentially from those of the received first-order logic only because tertium non datur sometimes fails. We therefore have to look out for such failures. What are they like? For one thing, (S ∨ ¬S) is not any longer always true. The conditional (A ⊃ B) is still equivalent with (¬A ∨ B), but what it now says is that if A is true or indefinite, B is true.

Now we have in fact reached a solution of the bald man paradox. What has been shown is that the paradox should be discussed in terms of IF logic rather than traditional
first-order logic. If so, the inference from (P.1)* and (P.2)* to (C)* is indeed valid. But the second premise is not true. For what it now says is that if a man with n hairs is bald or neither-bald-nor-not-bald, then a man with n+1 hairs is bald. This is obviously not true.

This result is as full a solution to the sorites paradox as one can hope. Not only does it show how to avoid the paradox in our formal discourse. It offers an educated diagnosis of how the paradox comes about. In a sense, it vindicates the truth-value gaps idea. For the only substantial assumption we had to make is the failure of the law of excluded middle which is a generalization of the truth-value-gaps assumption. And the failure of the law of excluded middle is made inevitable by IF logic which in turn is indispensable for our logical language to serve adequately its task. The evidence for the unavoidability of IF logic is independent of the sorites paradox. In this sense, we have solved the paradox without any additional assumptions beyond simply the meaning of logical constants which is built into the interpretation of the underlying logic. In other words, the paradox has been solved by purely logical means without evoking any pragmatic or other extralogical ideas. The solution can be considered as being merely a consistent way of implementing the initial hunch that the source of the paradox is the existence of “truth-value gaps”, that is, of cases where the law of excluded middle fails for the crucial predicate. The reason why earlier attempts to carry out this idea led to further problems is that in the absence of IF logic these attempts were not coherent. The tertium non datur was not allowed to fail in all cases.

Does this solve the sorites paradox for good? Can editors of philosophy journals now safely reject all new papers on the sorites paradox? Yes and no. The solution
reached here is definitive, but it is not the whole story. It seems to concern only
languages whose whole logic is captured by IF logic. But in natural language and indeed
in every adequately strong language the contradictory negation is also present. The
discovery of IF logic shows a remarkable thing about these languages. (See here Hintikka
2002(a).) Since the classical semantical rules that are obviously followed in our native
logic give rise to a negation different from the contradictory one, it must be present in
any normal language. Hence, *in all sufficiently strong languages there are inevitably two
different negations present*, the strong (dual, game-theoretical) negation and the
contradictory negation. This is the case also in natural languages even if there usually is
only one kind of negation lexically and syntactically present. This insight throws in fact
an interesting light on several phenomena in the grammar of negation in natural
languages. (For such grammar, see e.g. Horn 1989.)

Thus our examination of the sorites paradox needs a second stage because one can
try to formulate the paradox by means of a logic that goes beyond IF logic. For one thing
we have to consider (and use) richer languages in which the two negations are present.
Why does the sorites paradox not arise in them? Or does it?

The answer to this question obviously depends on a preliminary one. How can
the contradictory negation be introduced into a purely IF language? Since the classical
semantical rules give rise to the strong negation, we can without changing our semantics
introduce the contradictory negation $\neg S$ (as distinguished from the strong negation $\sim S$)
only by stipulating that $\neg S$ is true if and only if $S$ is *not* true. Here the stress indicates, as
it often does in natural language, contradictory negation in the metalanguage. It follows that \( \neg \) can only occur sentence-initially but not prefixed to open formulas.

The resulting logic has been called extended IF logic. It can be considered our true basic logic. It is elementary in the sense that in its interpretation we need no applications of tertium non datur to closed totalities and in the sense that in its applications (in the relevant “language games”) no infinite moves are needed. Extended IF logic is symmetrical between its two halves. The original IF half is equivalent with the \( \Sigma^1_1 \) fragment of second-order logic while its mirror image half (“the \( \neg \) half”) is equivalent to the \( \Pi^1_1 \) fragment. In IF logic there exists a complete disproof procedure but no complete proof procedure, while in the negated half the converse situation prevails. Each half admits of a truth predicate formulated in the same half. Algebraically, the extended IF logic exhibits a well-known structure: it is a Boolean algebra with an operator. (See Hintikka 2004.)

The question now becomes: Does the sorites paradox rear its head in extended IF logic? In order for it to arise, we would have to use a conditional in (P.2)** which would not make them equivalent.

(3) \[ \neg A(n) \lor A(n+1) \]

but rather to

(4) \[ \neg A(n) \lor A(n+1) \]
(Here \( n \) is a free numerical variable.) Similar things can be pointed out \( \textit{mutatis mutandis} \) of (P.2)*.

But in (4) the contradictory negation \( \neg \) occurs prefixed to an open formula. (In a negation normal form of the principle of induction, \( \neg \) would occur within the scope of an existential quantifier.) This is not allowed in the extended IF logic. Hence the paradox of the bald man does not arise in it, either.

Here we can in fact have a closer view of the sources of the paradox. In the case of the bald man version, it can trivially be assumed that

\[
\neg A(n+1) \lor A(n)
\]

Together with (4), this implies

\[
\neg A(n) \lor A(n)
\]

In other words, the principle of mathematical induction does not produce the paradox unless the predicate (simple or complex) figuring in it obeys the law of excluded middle. This condition is not satisfied in the bald man case. In (P.2)** the (usually complex) operative predicate is \( A(x) \) while in (P.2)* it is \( (H(x,n) \supset B(x)) \) which has truth-value gaps as soon as \( B(x) \) does.
What makes formulas like (1)*-(2)* paradoxical is the fact that in them ordinary first-order logic with its law of excluded middle is used in the teeth of truth-value gaps. The same holds of the usual form of mathematical induction.

We can thus express one important upshot of these observations in general terms by saying that the principle of mathematical induction is applicable (without additional assumptions) only if the law of excluded middle applies to the operative predicate. Thus the problematic of the sorites paradox is closely connected with the principle of mathematical induction. No wonder that the two have frequently been discussed together.

Because of this close connection, it is in order to make a third round of comments on negation and on mathematical induction. They lead us to yet new logical languages, languages beyond an IF language extended by sentence-initial contradictory negation. It is in fact possible to give a meaning to the contradictory negation \( \neg \) also when it is used within the scope of a quantifier. However, doing that requires going beyond extended IF logic and beyond game-theoretical semantics. Consider, for the purpose of seeing this, the following pair of sentences

\[
(7) \quad (\forall x)(A(x) \& \neg(\forall y)A(y))
\]

\[
(8) \quad (\forall x)(\neg\neg A(x) \& \neg(\forall y)A(y))
\]
Either of them can be taken to represent the logical form of the negation of (1).

However, game-theoretical semantics does not assign an interpretation to (8). There nevertheless is an obvious way of doing so. What we can do is to say that (8) is true if and only if for each member $m$ of the domain of individuals, the verifier has a winning strategy in the game connected with

$$(9) \quad \neg \neg A(m) \& \neg (\forall y)A(y)$$

In (9), the contradictory negation $\neg \neg$ occurs only sentence-initially, since $m$ is a constant. Hence it can be interpreted game-theoretically. The existence of a winning strategy for the verifier for all values of $m$ can indeed serve as a truth-condition of (9). But then (8) can be taken to be true because (9) is true for all values of $m$. For $\neg \neg A(m)$ says merely that $A(m)$ is not false. It can still fail to be true, making $(\forall x)A(x)$ true. Hence (1) can be false, and hence it is not logically true. This reflects the fact that (9) is true for all values of $m$ only if $A(x)$ satisfies the law of excluded middle.

The extended truth condition illustrated by this example is nonelementary in a way the truth condition for IF logic is not, for it relies heavily considering the domain of discourse (which may be infinite) as a closed totality. However, if we are not worried about this nonelementary (“infinitary”) character of the resulting logic, we can allow (starting from extended IF logic and hence moving entirely on the first-order level)
arbitrary nesting of the contradictory negation. The result is a first-order logic that is as strong as the entire received second-order logic. (This is shown in Hintikka 2006.)

In the resulting logic, the original formal derivation of the paradox would be valid. In other words, in the new, much stronger language (unlike the two languages considered earlier, viz. unextended and extended IF languages) there is a valid formalized version of the argument that leads to the paradox. It is (P.1)*- (P.3)* with \( A \supset B \) interpreted as \( \neg A \lor B \). Does that mean that we still have a paradox in our hands? No, because the only thing that can now be concluded is that the formal argument would no longer be a faithful representation of the ordinary language argument from (P.1)-(P.2) to (C). Since in (P.2)* and in (P.2)** the conditional occurs within the scope of a universal quantifier, from the point of view of traditional first-order logic their semantics is that of second-order logic. The validity of the formal derivation of the paradox belongs to the same category of puzzles as the validity of the well-known first-order formulas that seem to express ordinary-language sentences that are not valid intuitively at all.

We have encountered cases in point in the form of (1)*-(2)*. Their status as logical truths turns on the interpreting of \( A \supset B \) as \( \neg A \lor B \), which makes (1)*-(2)* wrong representations as the logical form of the ordinary language statements (1)-(2) when truth-value gaps are present, as they implicitly are in (1)-(2). Now we have seen that from the point of view of the fully extended IF logic the paradoxicality of the bald man paradox is of the same kind as that of (1)*-(2)*.
In terms of the reconstruction of second-order logic on the first-order level, we can thus give an interpretation to the principle of mathematical induction even the second premise (P.2)** as understood as having the form

\[(10) \quad (\forall n)(\neg A(n) \lor A(n+1))\]

However, the principle of mathematical induction is still on this construal applicable only when the predicate \(A(x)\) obeys the law of excluded middle, that is, does not have any truth-value gaps. For if it does, then it could be false in that for some \(n\) \(A(n)\) might be neither true nor false (and hence make \(\neg A(n)\) and (10) true) while \(A(n+1)\) is false. Thus the role of mathematical induction in the resolution of the paradox remains the same.

The conclusions we have reached concerning mathematical induction might not at first sight affect the uses of mathematical induction in mathematics, for the nonlogical predicates and functions used there do not usually admit truth-value gaps. But a complex predicate may fail to obey the law of excluded middle also because of the presence of informational independencies (expressed by the slash). Now IF logic with its slash notation has not been explicitly used in ordinary mathematical research. However, it has been pointed out repeatedly (see e.g. Hodges 2006) that informational independence and hence IF logic has tacitly been used in actual mathematical reasoning. Hence one can be sure of the applicability of mathematical induction to some complicated special case only by first analyzing the propositions involved logically so as to ascertain that no hidden
informational independencies lurk somewhere in them. Otherwise an actual working mathematician might end up as a victim of the sorites paradox. Whether such fallacies are actually found in the mathematical literature requires a wider search of possible examples and a closer examination of them than can be undertake here. This sorites paradox is hence not merely philosophical logicians’ self-inflicted problem. The bald man calls our attention to a significant issue in the actual proof methods of mathematics.

What has been found here has also more philosophical implications than can be explored in a single paper. For instance, the fact that mathematical induction does not work for concepts with truth-value gaps provides a perspective on the thought of those philosophers who have emphasized the unsharpness of most of our actual concepts, for instance Wittgenstein with his ‘family resemblance’ idea. Such philosophers can be expected to have doubts about mathematical induction, too.

In particular, it has been claimed that the presence of unsharp concepts in our language means that it does not have a definite logical structure, a structure that can be captured by an explicit formalization. This claim has been refuted here. We have even seen what the main vehicle of the formalization of unsharp concepts is: IF logic. This puts the entire theory of vagueness into a new light. Is the bald man perhaps telling us that IF logic is among other things the true “fuzzy logic”? 
Literature

Feferman, Solomon, 1996, “Kreisel’s ‘unwinding’ program,” in Kreiserliana ed. by
(evenly sec. 2)


Haufling, Oswald, 2000, “What is wrong with sorites arguments”, Analysis vol. 61, pp. 6.


Bulletin of Symbolic Logic vol. 4, pp. 303-337.


Hodges, Wilfred, 2006, “The logic of quantifiers”, in The Philosophy of Jaakko Hintikka

(Library of Living Philosophers), ed. by Randall Auxier, Open

Court, Chicago, pp. 521-534

Horgan, Terry, 1990, “Psychologistic semantics, robust vagueness, and the philosophy of

language,” in Belief Systems in Language: Studies in Linguistic Prototypes, ed. by


Horn, Laurence R., 1989, A Natural History of Negation, The University of Chicago

Press, Chicago.

Keefe, Rosanna, and Peter Smith, editors, 1987, Vagueness: A Reader, MIT Press.

Lakoff, George, 1975, “Hedges: A study in meaning criteria and the logic of fuzzy

concepts,” in Contemporary Research in Philosophical Logic and Linguistic


