Epistemic Paradoxes

I. The Surprise Exam Paradox
The teacher announces "at some point this week there will be a surprise exam." The students reason as follows: it can't be that the surprise exam will be on Friday, because then after class on Thursday, we'd know when the exam is. And then it wouldn't be a surprise. So the last day the exam could be on is Thursday.

	Monday	Tuesday	Wednesday	Thursday	Friday
Exam?	Maybe	Maybe	Maybe	Maybe	No

But it can't be on Thursday, because then, after class on Wednesday, we'd know that it was on Thursday and it wouldn't be a surprise. So the last day the exam could be on is Wednesday.

	Monday	Tuesday	Wednesday	Thursday	Friday
Exam?	Maybe	Maybe	Maybe	No	No

But it can't be on Wednesday, because then, after class on Tuesday, we'd know when the exam is. And then it wouldn't be a surprise.

	Monday	Tuesday	Wednesday	Thursday	Friday
Exam?	Maybe	Maybe	No	No	No

Etc… So there can't be a surprise exam.

II. The Lottery Paradox
Here's a principle that looks good. If you rationally believe the premises of a valid argument, it's rational to believe the conclusion of that argument.
Here's a valid argument:
P1
P2

P1 & P2
(If you believe P1 and you believe P2, you should also believe P1 & P2. Similarly, if you believe P1 and you believe P2….and you believe Pn, you should also believe P1 & P2 & …Pn.
However, here's a seeming counterexample:

Jones has a lottery ticket and there is a 1 in a billion chance that his ticket will win. Quite sensibly, he believes that his ticket won't win. His friend Smith also has a ticket and Jones also quite sensibly believes that Smith's ticket won't win. Similarly, for any ticket i of the billion tickets, Smith believes that ticket i won't win. Let's call the proposition that ticket i won't win: $\sim t_i$. Jones believes $\sim t_1$, and Jones believes $\sim t_2$, and Jones believes $\sim t_3$….and Jones believes $\sim t_{1\ billion}$. But now we can construct the following valid argument:

$\sim t_1$
$\sim t_2$
$\sim t_3$
.....
$\sim t_{1 billion}$
___
$\sim t_1$ & $\sim t_2$ & $\sim t_3$...... & $\sim t_{1\ billion}$.

By the principle above, since Jones believes all of the premises of this argument, Jones should believe the conclusion. In other words, Jones should believe that no ticket will win! But surely that's wrong!

III. The Dogmatism Paradox
Definition: E is misleading evidence against p, if E is evidence against p and p is true.

Suppose you know that p is true. Then you know that any evidence against p is misleading (since it is evidence against a true proposition!). So you should ignore any evidence you get against p.
But of course that's crazy!

Here's the valid argument
P
P→(if there's evidence against p – the evidence is misleading)
___
If there's evidence against p – the evidence is misleading

Problems of Philosophy
Epistemic Paradoxes - Solutions

I. The Surprise Exam Paradox Solution (?)
What does it mean for the exam to be a surprise? Let's suppose the following:

An exam on day D is a surprise if and only if, at no time before D, do the students know that the exam will be on D.

The paradox relies on the assumption that the exam can't be on Friday, because if it were, it wouldn't be a surprise. This means that if the exam were on Friday, at some point before Friday, the students would know that the exam would be on Friday. We supposed, they'd know on Thursday when the exam is. But why is this true? Let's consider each possibility: a) on Thursday, they'd know that the exam will be on Friday, b) on Thursday, they wouldn't know that the exam will be on Friday.

   a)   Suppose that on Thursday they'd know that the exam will be on Friday. Then, if they'd know that they know that the exam will be on Friday, then they'd know that the exam won't be a surprise.   But if they know the exam won't be a surprise, then they don't know on Thursday that the surprise exam will be on Friday. So they *don't* know on Thursday that the exam will be on Friday.

But….

    b) Suppose that on Thursday they wouldn't know that the exam will be on Friday. Then, if they'd know that they don't know that the exam will be on Friday, then they'd know that if there's an exam on Friday, it will be a surprise. So it's possible for there to be a surprise exam on Friday. But since they know there will be a surprise exam sometime this week, and it could be Friday, and it hasn't been any other day, then they *do* know that the exam will be on Friday. But..(see (a))

Anyway, it's not completely straightforward that the claim "on Thursday, they would know that the exam will be on Friday" and this is what's necessary to get the paradox going. Maybe the right thing to say is that they *wouldn't* know on Thursday that the exam will be on Friday, because at that point they should assume that the teacher has forgotten (or something like that). They shouldn't then conclude from the fact that they don't know that there will be an exam that, in fact, there will be an exam because it's possible for them to be surprised.

II. The Lottery Paradox Solution

Some people take this paradox to motivate using degrees of belief instead of on/off belief.

Degrees of belief are measured by real numbers between 0 and 1 called credences where having credence 1 in p is certainty in p, and 0 in p is certainty in ~p. Let's simplify things so that there are only 100 tickets in the lottery. Jones' credence that he will win is 0.01. For any ticket i, Jones' credence that ticket i will win is 0.01. However, while it may be true that if you *believe* all of the premises of the valid argument you should believe the conclusion, it's not true that if you *have high credence* in all of the premises of a valid argument, you should have high credence in the conclusion.

So Jones has the following perfectly reasonable credences

$Pr(\sim t_1) = .99$

$Pr(\sim t_2) = .99$

$Pr(\sim t_3) = .99$

….

$Pr(\sim t_{100}) = .99$

$Pr(\sim t_1 \,\&\, \sim t_2 \,\&\, \sim t_3 \,\&\, \sim t_4 \,\&\ldots \sim t_{100}) = 0$

But…even if we can make sense of what's going on when we talk about credences, we frequently talk about beliefs, so how do we resolve the paradox put in terms of beliefs? Perhaps what we should say is that Jones shouldn't *believe* he won't win the lottery. (A motivation for this: if he did believe he wouldn't win, shouldn't he just throw away the ticket?).   Note that this rules out certain views about how degrees of belief are connected to on/off belief.

III. The Dogmatism Paradox - Solution

First Pass: It's true that you can *now* rationally believe that any evidence against p is misleading. But if you were to get evidence against p, you would no longer rationally believe p. So you wouldn't be able to rationally conclude that evidence against p is misleading, and so you shouldn't disregard the evidence.

But…even if it's true that you shouldn't disregard the evidence when you get it, if you rationally believe now that any evidence against p is misleading, shouldn't you try to avoid getting evidence against p?

Second Pass: This paradox too is more easily solved by moving to a probabilistic framework. Here's what we've got:

Pr(P) = high
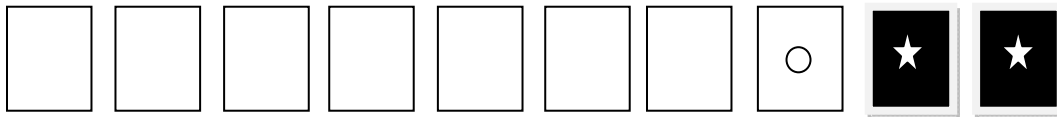Pr(P → (if there's evidence against p -> the evidence is misleading) = high
But…
Pr(if there's evidence against p → the evidence is misleading) = low

We already saw in the lottery paradox that just because you have credence in the premises of a valid argument doesn't mean you should have high credence in the conclusion. In this case, what's going on is this: P is probably true. And because of this you probably won't get evidence against P. But if you do get evidence against p, it's probably not misleading.

Example: There are 10 cards. 8 of them are white and two of them are black. Most of the white cards are plain but one has a circle on it. The two black cards both have a diamond on them. You pick a card at random.

Pr(card is white) = high (.8), Pr (card is white → (if it has a picture on it, it is a circle) = high (1)
Pr (if it has a picture on it, it is a circle) = low (.333)
Similarly, let's consider a bunch of possible worlds. Let E be evidence against p and "ml" indicate that the evidence is misleading.
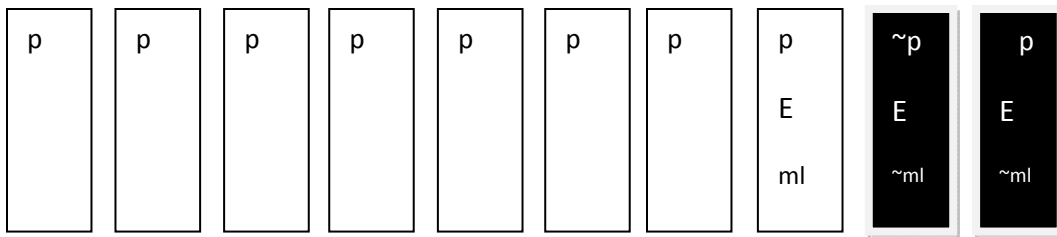
So: Pr(p) = 8, Pr(p→if there's evidence against p, its misleading) = 1, but Pr(if there's evidence against p, its misleading) = .333

24.00 Problems in Philosophy
Fall 2010