Paradox Philosophy

Course at Akita International University

Philipp Blum, University of Lucerne, philipp.blum@philosophie.ch
https://philipp.philosophie.ch/teaching/paradoxes20.html

Summary of the course, version of February 17, 2020

Contents

1 Administrative info

Course description. Why is it that some things are too big to exist? Why can't things not include themselves or be too big to be thought about? Does Achilles ever overtake the tortoise, is motion even possible? Does the liar who says "I'm lying" even say something; if so, is it true what he says? If everyone does what is best for him no matter what the others do – will then all be better off? if I were offered a blessed life in the Matrix, filled with only the experiences I cherish most, should I take the offer? is there anything valuable at all?
These are interesting, but also difficult question. In this beginners' course, students will become acquainted with some ways of asking them and be encouraged to think about answers.
In this course, we will encounter five different paradoxes, hitherto unsolved, from different areas of philosophy. Attending students will get brief presentations beforehand, will learn about the background assumptions and the theoretical context, try their hands at solutions (group work of about 30 minutes) and then be presented with an overview of extant diagnoses, being given the opportunity to critically discuss them.

Pitch. You will learn that some sets are not groups and some groups are not sets, that Achilles cannot overtake the tortoise, that, indeed, there is no motion at all and everything is one, that some things cannot be said, nor believed, nor known, are true and false, that you can prove that the moon is made of cheese, that no-one and everyone is bald, that there cannot be surprise exams, that you should assign a probability of 1/3 to a fair coin falling heads, that there are intentions you cannot form, beliefs you cannot assert, but that you can see the back side of the moon.

Target audience. The course is designed for students who are ambitious enough to test their critical thinking skills at some of the most difficult problems at all. They will benefit first, by learning humility, that nothing is quite as easy as it first appears, and second, by appreciating their own intellectual progress as they revise, change, refine and develop a first response to the problems, which are likely to accompany them for the rest of their lives. An important task is to justify and defend their positions; students will thus also improve their debating skills.

Schedule. We will devote each day to a different family of paradoxes:

Monday Paradoxes of size – the set-theoretic paradoxes (Cantor, Burali-Forti, Russell, Prior-Kaplan)

Wednesday Paradoxes of motion – Zeno, but also more contemporary versions (0-valued quantities, instantaneous velocity, in-/finite time)

Thursday Paradoxes of truth – the Liar, semantic paradoxes (Grelling, Curry)

Friday Paradoxes of decision-theory and knowledge – Prisoners' Dilemma, Newcomb, the unexpected examination

Monday Paradoxes of value - the experience machine, the toxin puzzle, bootstrapping

To get a mark for this course, do the following:

- Choose a paradox, either from the course material, or from Clarke or from Sainsbury.
- Explain it briefly in your own words.
- Either explain what is paradoxical about it, in your own words, and briefly present a solution to it (of course, the 'solution' does not have to be your own, you can pick one that others have suggested).
- Critically discuss this solution and raise some problems about it.
- Save your essay as a pdf file and send it to philipp.blum@philosophie.ch before the 21st of February.

2 What is a paradox?

Paradoxes have a long history of both provoking and frustrating thought, often challenging unproved assumptions and uncovering surprising equivocations. Sainsbury (1995, 1) characterises them as follows:

> This is what I understand by a paradox: an apparently unacceptable conclusion derived by apparently acceptable reasoning from apparently acceptable premises.

Similarly, Quine (1962, 1) says:

> …a paradox is just any conclusion that at first sounds absurd but has an argument to sustain it.

Generalising a little, Schiffer (2003, 68) and Lycan (2010, 618) call a paradox an inconsistent set of propositions which are all plausible.

But is a paradox just a puzzle? I would think a good, interesting paradox is more. Some 'paradoxes' (scare quotes) are just fallacies, some are real dilemmas, some perhaps even antinomies. Horwich (2016, 213) gives a better characterisation: "a philosophical paradox is a battery of a priori considerations that engenders conflicting epistemic inclinations".

Paradoxes are useful tools of philosophising because they narrow down the dialectic possibilities to three:

- argue that the conclusion is not as unacceptable as it appears, that we can 'live with it';
- argue that the reasoning is not as convincing as it appears, perhaps by uncovering a hidden ambiguity in one of the key terms;
- argue that the paradox is a reduction to the absurd (a 'reductio', for short) of one of the premises and motivate independently its rejection.

More generally, a paradox presents us with two jobs:

- the diagnostic job is to 'deal with' the paradox in one of the three ways above and, importantly, to explain why the reasoning struck us as a paradox in the first place – it's not enough, e.g., just to reject one of the premises as false; we also have to explain why it appeared true to us (or to some people more intelligent than we are);
- the explanatory job is to develop a positive theory of the subject-matter, which does not 'fall prey' to this or similar paradoxes.

Schiffer (2003, 69) argues that even if the first job cannot be done, we may still be able to provide a "unhappy-face" solution, by showing that and why it cannot be done. This, however, equivocates on "cannot" – that some paradoxes have hitherto resisted 'resolution' should not lead us to expect that they cannot, in principle and forever, be solved, nor that no progress in this direction can be made.

3 Mathematical Paradoxes

3.1 Things and their groupings

Things group together in various ways: many different little parts compose a whole, some notes a tune, a group of people a family. Some of the properties of the groupers are inherited by the group, some not: if the little parts are out of wood, the whole is, at least partly, wooden, but the note may be harmonious while the tune is not, the family may be five while each member of it is just one.

Sometimes, groups have the same properties as their members in the same way, sometimes they have the same properties, but in different ways. The group of large things is itself large, but in another way; the group of all the things I like is something I like, but in another way. Not so with other properties: the group of things in Japan is in Japan, the group of precious things is itself precious, the group of all groups that were mentioned here today is itself a group.

Groups do not just have properties, and sometimes the same properties as their members, they can also be defined by them. I can characterise the very same group in two different ways, called 'extensionally', by enumeration, or 'intensionally', by a condition all and only its members fulfil:

$$X := \{\text{Huey}, \text{Dewey}, \text{Louie}\} \qquad X := \{x \mid x \text{ is a child of Donald Duck}\}$$

For many groups, I need to impose further conditions: for some things to be a *family*, for example, it is not enough that they exist; they also have to be (in some legal sense) relatives of each other. For other things, more 'material' relations are needed: for a number of wooden planks to constitute a ship, for example, they have to be in contact with each other and together be arranged in a certain form.

For the 'lightest' groups, what mathematicians call "sets", no such conditions are required: the mere existence of the members guarantees the existence of the set – or so it seems at least!

In the case of sets, I have defined the very same set of exactly three members above, extensionally and by enumeration on the left, intensionally and by a condition on the right.

Though this is not required in all cases,[1] several phenomena suggest treating 'groups' as new entities, something 'over and above' their members:

- While the members are many, the group is one.
- In many cases, the group is somewhat more abstract, more aloof, more 'formal' than its members.
- In many cases, the group seems to have different properties than its members.

Once we make this step for sets, however, paradox looms.

3.2 Cantor and modern set theory

The German mathematician Georg Cantor was interested in the use of infinities in mathematics. The infinity of the natural numbers puzzled already the Greeks:

$$\mathbb{N} := \{1, 2, 3, 4, 5, 6, \dots\} \qquad \mathbb{N} := \{x \mid x \text{ is a natural number}\}$$

1. For many – some say: for all – 'ordinary' uses, talk about groups can be 'paraphrased away', translated into talk not about groups but just plurally about several things simultaneously. "Donald Duck takes the group of his children to the movies", e.g., is just a funny way of saying the same thing as does "Donald Duck takes his children (or: Huey, Dewey and Louie) to the movies".

How 'big', in an intuitive sense, is this group? Is it bigger, for example, than the set of even natural numbers:

$$\mathbb{E} := \{2, 4, 6, 8, 10, 12, \dots\} \qquad \mathbb{E} := \{x \mid x \text{ is a natural number} \wedge x\, x\text{is even}\}$$

The answer may seem simple: *obviously*, $\mathbb{N}$ is bigger than $\mathbb{E}$!

- First, all members of $\mathbb{E}$ are members of $\mathbb{N}$, but not all members of $\mathbb{N}$ are members of $\mathbb{E}$: for example, 1, 3, 5 are not!
- Second, the set $\mathbb{E}$ is defined by an extra clause ("being even") which is not true of all the members of $\mathbb{N}$: to its a real condition and it is thus 'harder' to be a member of $\mathbb{E}$ than it is to be a member of $\mathbb{N}$ ("it's not enough to be a natural number, you also have to be divisible by 2…").

It was therefore surprising that the German mathematician Cantor defended the opposite answer: in the mathematically relevant sense, both sets are equally big – they have the same *cardinality*. They have the same cardinality because there is a one-to-one correspondence (mathematically: a bijective function) between their members, namely multiplication by 2.

- for every member of $\mathbb{N}$ there is exactly one member of $\mathbb{E}$, which is the double of that first member of $\mathbb{N}$;
- for every member of $\mathbb{E}$ there is exactly one member of $\mathbb{N}$, which is the half of that first member of $\mathbb{E}$.

Not only does the set of natural numbers have the same cardinality than the set of even numbers, it even has the same cardinality as the set $\mathbb{Q}$ of rational numbers! But not, Cantor showed by his famous 'diagonalisation' argument, as the set $\mathbb{R}$ of real numbers. If we assume the real numbers being numbered, arranging them in an infinite list, we can define a number which differs from any number in the list by at least one digit – this number is in $\mathbb{R}$ but not on the list, of which it follows that there are non-denumerably many real numbers.

As the usefulness of sets became apparent and set-theory became the foundation of present-day mathematics, a first paradox ensued: what about all sets – do they form a set? Intuitively, if we know what a set is (there's not much to it, it's just a bare collection), we can define the set of all sets as follows:

$$\mathbf{V} := \{\mathbb{N}, \mathbb{R}, \mathbb{E}, \text{the set of all cows}, \{\text{Huey}, \text{Dewey}, \text{Louie}\}, \dots\}$$

or, by a condition, as follows:

$$\mathbf{V} := \{x \mid x \text{ is a set}\}$$

If we do so, however, we can ask: is V a set?

- If it is, then it is a member of V (because that's how V was defined): but then $V \in V$, i.e. the set contains itself – but does it then also contain other sets, such as the set of all cows? what is left of it, if we strip it of just one of its (many) members, namely of V?
- But if it is not, then not everything that is definable as above, not every collection of things is a set – what further conditions are required?

Worse was to come: the so-called 'diagonalisation argument' Cantor gave to show that the cardinality of $\mathbb{R}$ was strictly greater than that of $\mathbb{N}$ (mathematically: that $\mathbb{N}$ can be embedded into $\mathbb{R}$ but that there is no one-to-one correspondence of $\mathbb{R}$ to $\mathbb{N}$) can be generalised to Cantor's theorem: that there are always more subsets than members (Cantor, 1892). If we start with a set X and form the set of its subsets (the so-called "powerset of X", $\mathcal{P}(X)$), we get a set of higher cardinality than X. If there were a set V of all sets, then we would have $|\mathcal{P}(V)| \leq |V|$, i.e. the cardinality of its powerset would not be greater than that of the set V itself (for all the subsets of V are also sets), which contradicts Cantor's theorem.

3.3 Frege's definition of number and Russell's paradox

Despite its unapplicability to V, the concept of cardinality proved very fruitful. The mathematician and philosopher Gottlob Frege, arguably the grandfather of modern Western philosophy, used it to define the concept of natural number (Frege, 1884). In the same way in which directions are what parallel lines have in common, Frege said, numbers are what equinumerous sets (ie. sets of the same cardinality) have in common. By identifying sets, in effect, with conditions, Frege wanted to provide a purely logical foundation for the whole of mathematics.

Unfortunately, this project foundered. It was in the first volume of Frege's major work, "The Fundamental Laws of Arithmetic" (Frege, 1893), that the English philosopher Bertrand Russell discovered a contradiction. Russell showed that Frege's rules permitted the definition of a set, the existence of which allowed the derivation of a contradiction from the system and thus showed the system to be inconsistent (i.e.: useless).

The set in question was the set of all sets that do not contain themselves. Some sets do not contain themselves, for example the set of cows. It's a set, and not a cow, and to be a member of it something has to be a cow, so it is not a member of itself. There are many other such sets: the set of people, the set of colours, the set of Donald Duck's childrens and so on. So let us define the set of them and call it "Hugo":

$$\mathbf{H} \ := \ \{x \mid x \notin x\}$$

Now we know what Hugo is and we know that Hugo is a set, so we can ask a simple question: does Hugo, or does it not, contain itself? The problem is that neither answer is possible:

- yes, that is: $H \in H$. But then Hugo has to satisfy the condition that defines it, which is: *not containing itself*; so: if Hugo contains itself, it does not contain itself;
- no, that is: $H \notin H$. But then Hugo *does* satisfy the condition that defines it, and so it *does* contain itself; so: if Hugo does not contain itself, it contains itself.

This means that we are in a very uncomfortable situation. The problem is not that one possible answer is 'unstable' and leads to its negation. This is perfectly acceptable (it's called a proof 'by reduction to absurdity'): if the answer 'yes' to a yes/no question leads to the answer 'no', then we know that the answer is 'no'.

The problem is rather that *both* answers to the question are 'unstable' in this way; and that because the question is a yes/no-question, no 'middle ground' is possible. We have obtained a proof of the following claim:

$$\text{Hugo} \in \text{Hugo} \iff \text{Hugo} \notin \text{Hugo}$$

This is a straight contradiction, a reduction to absurdity of some of the assumptions that have gone into its derivation. The only such assumption was a completely innocuous one – that there is a set of all the sets that do not contain themselves.

We *have to* conclude that there is no such set. But it is difficult to see why not. To see this, compare the case with the barber who shaves all and only those people of his hometown who do not shave themselves. Because he would shave himself if and only if he does not, there cannot be such a barber. *Tant pis*, as we had no antecedent reason to believe in his existence anyway. With sets, the situation is different: to be able to talk meaningfully about the set of natural numbers or of Donald Duck's children, I had nothing more to do than to lay down their definition.

What more than the existence of its members *could* be required for the existence of a set?

Frege was devastated by Russell's derivation of a contradiction:

> Your discovery of the contradiction caused me the greatest surprise and, I would almost say, consternation, since it has shaken the basis on which I intended to build arithmetic, [...] In any case your discovery is very remarkable and will perhaps result in a great advance in logic, unwelcome as it may seem at first glance. (Frege, letter to Russell, 22 June 1902, quoted in van Heijenoort (1967, 127–128))

Of Frege's reaction, Russell said in a letter published for the first time in 1967:

> As I think about acts of integrity and grace, I realise that there is nothing in my knowledge to compare with Frege's dedication to truth. His entire life's work was on the verge of completion, much of his work had been ignored to the benefit of men infinitely less capable, his second volume was about to be published, and upon finding that his fundamental assumption was in error, he responded with intellectual pleasure clearly submerging any feelings of personal disappointment. (Russell to van Heijenoort, 23 November 1962, in van Heijenoort (1967, 127))

In most modern set-theories, not every condition defines a set. Sets are constructed axiomatically, as things that satisfy certain axioms. But we still do not know what they are.

4 Physical Paradoxes

4.1 Zeno's paradoxes

I'm following Salmon (2001a, 9–15) in the presentation of the paradoxes who in turn closely follows Russell's "lecture VI" of his 1914 (separately reprinted in Salmon (2001b, 45–58)), entitled "The Problem of Infinity Considered Historically".

1. Achilles and the tortoise.
    1. In order to overtake the tortoise, Achilles must run from his starting point to the tortoise's original starting point T_0.
    2. While he is doing this, the tortoise will have moved ahead to another point T_1.
    3. While Achilles is covering the distance from T_0 to T_1, the tortoise moves still farther to T_2.
    4. And so on.
    5. Whenever Achilles arrives at a point where the tortoise was, the tortoise has already moved a bit ahead.
    6. Hence, Achilles will never catch up with the tortoise.
2. The Dichotomy, first form.
    1. In order even to get to the end point of the course $T_0 = 1$, he must first cover the first half of the stretch to T_0, i.e. get to $T_1 = 1/2$.
    2. In order to get to T_0, he then has to cover half of the remaining distance, up to $T_2 = 3/4$.

   3. In order to get to T_0, he then has to cover half of the remaining distance, up to $T_2 = 7/8$.
   4. And so on.
   5. In order to get to T_0, Achilles has to complete an infinite number of runs.
   6. Hence, Achilles will never reach T_0.

2. The Dichotomy, second form.
   1. In order to complete the full distance to T_0, Achilles has to run the first half of it.
   2. In order to complete the first half of the distance to T_0, Achilles has to run a quarter of it.
   3. In order to complete the first quarter of the distance to T_0, Achilles has to run an eighth of it.
   4. And so on.
   5. Hence, Achilles cannot even get started.

These first two paradoxes question the idea that physical space is a continuum, i.e. infinitely divisible in ever smaller subspaces. The next two question the alternative view that physical space and time are atomistic, composed out of minimal parts.

3. The Arrow.
   1. Let a be an arrow in flight.
   2. At any given instant, a is where it is, occupying a portion of space equal to itself.
   3. During the instant a cannot move, for that would require the instant to have parts, because a would have to be in one place at one part of the instant, and in a different place at another part of the instant.
   4. (Moreover, for a to move during the instant would require that during the instant a must occupy a space larger than itself, for otherwise a has no room to move.)
   5. This holds of all instants.
   6. Hence, a is always at rest.

4. The Stadium. Consider three rows of objects, arranged in the following way:

$$\begin{array}{ccccc} & A_1 & A_2 & A_3 & \\ B_1 & B_2 & B_3 & & \\ & & C_1 & C_2 & C_3 \end{array}$$

While row A remains at rest, rows B and C move in opposite directions, until all three rows are lined up as shown here:

$$\begin{array}{ccc} A_1 & A_2 & A_3 \\ B_1 & B_2 & B_3 \\ C_1 & C_2 & C_3 \end{array}$$

In the process, C_1 passes twice as many B's as A's. Suppose space and time are atomistic in character (being composed of space-atoms and time-atoms of non-zero size). If the B's and the C's move at the same speed, ie. at the rate of one place per instant, then C_1 gets past B_2 without ever passing it: C_1 begins at the right of B_2 and ends up at the left of B_2, but there is no instant at which it lines up with B_2.

5. Plurality (elaborated version).
   1. If extended things exist, they must be composed of a plurality of parts.
   2. These parts are extended, so they must themselves have parts.
   3. And so on.

4. So there is an infinity of parts.
5. There must be some ultimate parts (parts that do not themselves have parts).
6. These ultimate parts cannot have a magnitude (because otherwise they could be further divided).
7. If the ultimate parts have no magnitude, then the object they compose cannot be extended.
8. If the ultimate parts do have a magnitude, then the object they compose must have an infinite magnitude.

(This is an imperfect version, as premise 5 is entirely unsupported – can anyone do better?)

4.2 Not easily solved by the modern calculus

Peirce was not impressed by the Achilles paradox:

> …this ridiculous little catch represents no difficulty at all to a mind adequately trained in mathematics and logic. (Peirce, 1935, 177)

It is true that contemporary mathematics stipulates that the sum of a convergent infinite series is the limit of the sequence of its partial sums.[2] But even if we know 'how to add' convergent series of numbers,[3] it is not clear that the mathematics of $\mathbb{R}$ is applicable to physical space, physical time or physical space-time.

Black (1951, 93) has offered a persuasive criticism of the claim that modern mathematics has 'solved' Zeno's paradoxes:

> [The mathematical 'solution'] tells us, correctly, when and where Achilles and the tortoise will meet, *if* they meet; but it fails to show that Zeno was wrong in claiming they *could* not meet. Let us be clear about what is meant by the assertion that the sum of the infinite series $100 + 10 + 1 + \frac{1}{10} + \frac{1}{100} + \ldots$ is $111\frac{1}{9}$. It does not mean, as the naive might suppose, that mathematicians have succeeded in adding together an infinite number of terms. […] To say that the sum of the series is $111\frac{1}{9}$ is to say that if enough terms of the series are taken, the difference between the sum of that finite number of terms and the number $111\frac{1}{9}$ becomes, and stays, as small as we please.

4.3 Whitehead's lesson: becoming is not continuous

Whitehead has seen in these paradoxes a proof that temporal processes are discontinuous:

> …if we admit that 'something becomes,', it is easy, by employing Zeno's method, to prove that there can be no continuity of becoming. There is a becoming of continuity, but no continuity of becoming. The actual occasions are the creatures which become and they constitute a continuously extensive world. In other words, extensiveness becomes, but 'becoming' is not itself extensive. (Whitehead, 1929, 53) / (Whitehead, 1978, 35)

In sct. II of ch. II of part II, Whitehead elaborates:

2. By 'stipulates' I mean the following: the limit of the partial sums, when it exists, is called the "sum" and abbreviated by "Σ", which are the same terms as those used for finite sums.

3. I put "how to add" in scare quotes because it has not been shown (and is not claimed) that this is an 'ordinary' case of addition.

> The argument [of Zeno], so far as it is valid, elicits a contradiction from the two premises: (i) that in a becoming something (*res vera*) becomes, and (ii) that every act of becoming is divisible into earlier and later sections which are themselves acts of becoming. Consider, for example, an act of becoming during one second. The act is divisible into two acts, one during the earlier half of the second, the other during the later half of the second. Thus that which becomes during the whole second presupposes that which becomes during the first half-second. Analogously, that which becomes during the first half-second presupposes that which becomes during the first quarter-second, and so on indefinitely. Thus if we consider the process of becoming up to the beginning of the second in question, and ask what then becomes, no answer can be given. For, whatever creature we indicate presupposes an earlier creature which became after the beginning of the second and antecedently to the indicated creature. Therefore there is nothing which becomes, so as to effect a transition into the second in question. (Whitehead, 1929, 106) / (Whitehead, 1978, 78)

Salmon (2001a, 18) paraphrases this as follows:

> An act of becoming is an indivisible unit; if you subdivide it in any way the resulting parts are not smaller acts of becoming.

Whitehead himself puts it this way:

> The conclusion is that in every act of becoming there is the becoming of something with temporal extension; but that the act itself is not extensive, in the sense that it is divisible into earlier and later acts of becoming which correspond to the extensive divisibility of what has become. (Whitehead, 1929, 107) / (Whitehead, 1978, 79)

But what is becoming then? How can something temporally extended come into being by an 'act' which is does not have temporal extension? Whitehead tries to explain this by saying that what comes into being is 'divisible' into different aspects ('prehensions') which 'concern' different parts of its temporal extension, but it is difficult to make sense of this.

Perhaps some light can be shed on this by comparing it to what Dowty (1978) and others have called the "imperfective paradox" (I prefer to call it a puzzle). Briefly, the puzzle is this: how can we characterize the meaning of a progressive sentences like "John was crossing the street" on the basis of the meaning of a simple sentence like "John crossed the street" when the first can be true of a history without the second ever being true (poor John is hit by a car)? Or, in a metaphysical pitch: how can something be a crossing and hence being 'directed towards' its completion without ever be completed?

4.4 Russell's lesson: the at-at theory of motion

While Whitehead takes Zeno to show the need for something like 'instantaneous' (temporally non-extended) becoming, Russell draws the opposite conclusion.

If we define with Cauchy the derivative as a limit rather than a ratio of infinitesimals and take motion to *be* (not just: to be mathematically represented by) a function from instants of time to points in space, we get a 'static' view of dynamic processes:

> Weierstrass, by strictly banishing all infinitesimals, has at last shown that we live in an unchanging world, and that the arrow, at every moment of its flight, is truly at

> rest. The only point where Zeno probably erred was in inferring (if he did infer) that, because there is no change, therefore the world must be in the same state at one time as at another. The consequence by no means follows …(Russell, 1903, 347)

Russell thus accepts Zeno's description of the arrow as forever at rest, but denies that this means that it is not moving over time: moving over time, according to Russell, just *is* being successively located at different places. Russell thus re-defines what counts as motion, a 'solution' strategy often employed in paradox philosophy.

Even independently of its conceptual inadequacy, the at-at theory is problematic:

- The at-at theory holds that the velocity of a moving body at an instant is grounded in its subsequent locations: the body has the velocity it has *because* it is at these different places at different times. According to an alternative theory, the explanatory relation holds in the other direction: the velocity explains, rather than is explained by, the different locations.
- It is unclear that the at-at theory affords us the resources needed to distinguish a rotating homogeneous disk from a stationary one.
- It does not distinguish in the right way between the state of a pendulum at rest and the state of a pendulum at one of the end-points of its swing. It distinguishes them in terms of velocity, while the real distinction should be in terms of momentum.
- It gives us a notion of instantaneous velocity which is extrinsic: the velocity an object has at an instant depends on where it is before and after that instant.

4.5 Black's lesson: the problem of hypertasks

Black saw the difficulty raised by Zeno's paradoxes as concerning the theory of action: does it make sense to suppose that someone has completed an infinite series of distinct runs? According to Black, it does not:

> The logical difficulty is that Achilles seems called upon to perform *an infinite series of tasks*; and it does not help to be told that the tasks become easier and easier, or need progressively less and less time in the doing. Achilles may get nearer to the place and time of his rendezvous, but his task remains just as hard, for he still has to perform what seem to be logically impossible. (Black, 1951, 94)

In support of Black, Thomson (1954) illustrates this impossibility with the example that has become known as 'Thomson's lamp': suppose a lamp is switched on and off an infinite number of times, during a finite time. At the end of that time: is it on or off? It cannot be on, because it was turned off again every time it was turned on; it cannot be off, because it was turned on again every time it was turned off. But it must be on or off, there is no other possibility.

If, however, an infinity of actions cannot be done, how can I ever do something? We're back at Zeno.

4.6 Background 1: the line and the points

In *Physics* VI.1, Aristotle concludes from a discussion of Zeno's paradoxes that the (real) line is not composed out of points. The details are murky: I understand that Aristotle thought the line is prior to the points, in the sense that the points are abstractions from the line. But why think either that this implies that the line is not composed out of the points? Why think the points are not parts of the line?

And can we tolerate such big a discrepancy between modern mathematics and the physical reality we use it to describe?

4.7 Background 2: the metaphysics of persistence and the problem of change

Things change: we grow older, acquire new properties and loose old ones, until we eventually die. This much is certain. The problem of change is to understand how this is possible. How can one and the same thing have incompatible properties? The problem, as it is the case with many others, may be alternatively put as an inconsistent set of intuitively plausible claims or as a paradox (plausible inference from plausible premisses to an implausible conclusion).

The inconsistent quatuor:

(i) If there is change, it is in respect to one and the same thing.
(ii) If there is change, it is in respect to one and the same property.
(iii) There is change.
(iv) Nothing both has and lacks the same property.

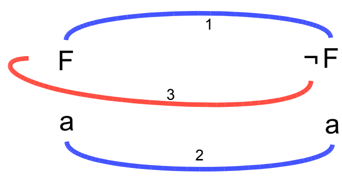
Denying each one of the premisses comes at a considerable price: denying (i) violates what they call "the proper subject condition", denying (ii) makes all change extrinsic, denying (iii) implies commitment to a certain kind of Parmenideanism, while denying (iv) violates the principle of non-contradiction.

The transformation into a paradox is immediate:

(i′) If there is change, it is in respect to one and the same thing.
(ii′) If there is change, it is in respect to one and the same property.
(iii′) Nothing both has and lacks the same property.
(iv′) Hence, nothing ever changes.

Change is problematic because it both seems real and impossible. It seems real because if there were no change, you could not convince anyone that there is no change, so we would all already believe it, which we don't. It seems impossible because to say that there is change is to say that some thing has a property [$\ldots_1$] and also lacks it [$\ldots_2$], which is a contradiction, hence not possibly true (and to understand why inserting "at one moment" for "$\ldots_1$" and "at another moment" for "$\ldots_2$" should do away with the contradiction just *is* the problem of change).

The ordinary concept of change seems to be internally inconsistent, or, at least, pulls us in different directions: for change, we need both constancy, of both objects and properties (blue lines), and variation, of being had and being lacked of the very same property by the very same thing (red line):



The problem of change is that the same pattern of sameness and difference is exhibited by the scenarios ruled out by the principle of non-contradiction, i.e. something's both having and not having the very same property.

According to Haslanger (2003), the puzzle of change arises from the inconsistency of the following five claims:

1. Persistence condition: Objects persist through change.
2. Incompatibility condition: The properties involved in a change are incompatible.
3. Law of non-contradiction: Nothing can have incompatible properties, i.e. nothing can be both P and not P.
4. Identity condition: if an object persists through change, then the object existing before change is one and the same object as the one existing after the change; that is, the original object continues to exist through change.
5. Proper subject condition: The object undergoing change is itself the proper subject of the properties involved in change.

Because there is change (1), objects persist (4) and the properties which they (5) have are incompatible (2) which is impossible (3).

The two main families of 'solutions' to the problem of change replace one of the blue lines above by a red one, denying either the incompatibility of the properties involved (2) or the numerical identity of the persisting object (4). The first view – endurantism – typically time-indexes the property: the thing is then said to 'change' from having F-at-t_1 to lacking F-at-t_2. This turns line 1 red: the two properties, *having* F*-at-*t_1 and *having* F*-at-*t_2 are not the same – what is had before the change is not what is lacked after it.[4] The second view either postulates temporal parts and attributes the properties to their whole (perdurantism: a is said to 'change' by having a temporal part that is F at t_1 and another temporal part that is not F at t_2) or postulates temporal parts and attributes the properties to them (exdurantism: a is said to 'change' by there being short-lived things a-at-t_1 and a-at-t_2 that are and are not F respectively). The problem with these solutions that none of them leaves room for change – which is why the occurrences of this word above are in scare quotes.

As in Russell, to be able to accept its possibility, the phenomenon itself is redefined.

5 Semantic Paradoxes

5.1 Truth and Falsity

According to the orthodox conception, every truth-apt sentence that is not true is false and every truth-apt sentence that is not false is true. This follows from the following three principles:

- The principle of **bivalence** which says that any truth-apt sentence is either true or false, i.e. that there is no third truth-value.
- The principle of **non-contradiction** which says that for no sentence both it and its negation are true, i.e. that if the sentence is true, its negation is not and vice versa.
- The principle of the **excluded middle** which says that for any truth-apt sentence, either it or its negation is true.

If we picture the logical space for truth-values as a square,

4. Or rather: if it is lacked after it, it is not because there was a change – *having* F*-at-*t_1 is also lacked, at t_2, by a thing that remained F throughout.

true	false

our three principles tell us to sort all truth-apt sentences and their negations into exactly one of the two halves.

- The principle of **bivalence** tells us that every sentence will be placed either in the left or in the right square;
- The principle of **non-contradiction** tells us that no sentence is placed in both squares at once;
- The principle of the **excluded middle** tells us that if we place a sentence into one of the squares, we have to place its negation in the other.

5.2 The Liar

Consider the following sentence and call it "the Liar":

(Liar) (Liar) is false.

By elementary reasoning, we derive a contradiction from the principle of bivalence (1):

1 (Liar) is either true or false.
2 If (Liar) is true, it is true that: (Liar) is false.
3 If (Liar) is true, (Liar) is false.
4 If (Liar) is false, it is false that: (Liar) is false.
5 If (Liar) is false, (Liar) is true.
6 (Liar) is true if and only if (Liar) is false.

2 and **4** substitute for the name of the sentence / statement / assertion the sentence / statement / assertion itself. This is acceptable if "it is true that" is a so-called extensional context.

The crucial steps are from **2** to **3** and, even more so, from **4** to **5**. The first step embodies one direction of the so-called "convention T", or principle of disquotation:

(T) It is true that p if and only if p.

By the left-to-right direction of (T), we move from "it is true that: (Liar) is false" to "(Liar) is false".

The step from **4** to **5** uses the principle of bivalence a second time, inferring truth from falsity of falsity. If we make the weaker assumption that falsity entails truth of the negation, we need another principle, the principle of the excluded middle to secure this step:

4 If (Liar) is false, it is false that: (Liar) is false.
4+ If (Liar) is false, it is true that: it is not the case that (Liar) is false.
4++ If (Liar) is false, it is true that: (Liar) is true.
5 If (Liar) is false, (Liar) is true.

Another, perhaps more direct way from to **5** is via the right-to-left direction of (T):

5− If (Liar) is false, it is true that: (Liar) is false.
5 If (Liar) is false, (Liar) is true.

The step from **5**− to **5** is intuitively motivated thus: "if (Liar) is false, then the world is as (Liar) says it is, hence (Liar) is true".

5.3 The Strengthened Liar

Could we treat the paradox as a reduction to absurdity of the principle of bivalence? One way to do so is to introduce a third 'truth value' ("indeterminate", "undecided", "shifty"). Could we not then just say that (Liar) is neither true nor false but 'shifty'?

Unfortunately, we cannot. Consider the so-called "Strengthened Liar":

(Strengthened Liar) (Strengthened Liar) is not true.

Whatever our third truth-value is, it certainly is incompatible with truth. This time, we derive a contradiction from an even more basic principle, the principle of non-contradiction:

1* (Strengthened Liar) is either true or not true.
2* If (Strengthened Liar) is true, it is true that: (Strengthened Liar) is not true.
3* If (Strengthened Liar) is true, (Strengthened Liar) is not true.
5−* If (Strengthened Liar) is not true, it is true that: (Strengthened Liar) is not true.
5* If (Strengthened Liar) is not true, (Strengthened Liar) is true.
6* (Strengthened Liar) is true if and only if (Strengthened Liar) is not true.

5.4 A Spicy Curry

A related, though different semantic paradox is Curry's:

(Sam) If (Sam) is true, the moon is made of cheese.

We derive an obviously false assertion about the constitution of the moon as follows:

1 Let us suppose that (Sam) is true.
2 (Under this supposition:) It is true that: if (Sam) is true, the moon is made of cheese.
3 (Under this supposition:) It is true that: (Sam) is true and if (Sam) is true, the moon is made of cheese.
4 (Under this supposition:) It is true that: the moon is made of cheese.
5 If (Sam) is true, the moon is made of cheese.
6 (Sam) is true.
7 The moon is made of cheese.

1 is surely innocuous: given that we understand the sentence, why should we not be able to suppose it to be true? The step to **2** is quite unproblematic too: we just substitute the sentence for its name. **3** appears equally unproblematic: certainly the supposition itself is true *under that supposition*!

The steps from **3** to **4** and from **5** and **6** to **7** are two applications of an inference rule logicians call "modus ponens". It's the rule that permits us to drop the antecedent of a so-called 'material conditional' once it has been established as true. This inference rule, within or outside of suppositional reasoning, is certainly unproblematic (indeed, 'blind' acceptance of it has been taken to be a hallmark of semantic competence with "if ...then"!).

The step from **4** to **5** discharges the supposition by conditionalising on it: if, under the supposition that p, I have shown that q, I can summarise my findings as "if p then q", the conditional not standing under any supposition.

The step from **5** to **6** again uses the right-to-left direction of the principle (T) above: **5** *is* what (Sam) says, so it implies its truth.

5.5 Having fun with liars and curries

Consider the 'paradox of the seducer', which consists in a series of two questions, which have to be answered by "yes" or by "no":

- Will you give the same answer to this and to the next question?
- May I have a kiss?

No matter what the answer is to first question, the answer to the second (arbitrary) question is "yes".

The appendix contains some 'problems' from Raymond Smullyan, who has written a lot of books full of funny logic puzzles:

"What is the Name of this Book? The Riddle of Dracula and Other Logical Puzzles", (1978)
"This Book Needs No Title. A Budget of Living Paradoxes", (1980)
"5000 B.C. and other Philosophical Fantasies", (1983)
"Lady or the Tiger? And Other Logic Puzzles Including a Mathematical Novel That Features Gödel's Great Discovery ", (1982b)
"Alice in Puzzle-Land. A Carrollian Tale for Children Under Eighty ", (1982a)
"Satan, Cantor, and Infinity and Other Mind-Boggling Puzzles ", (1992)
"The Riddle of Scheherazade: And Other Amazing Puzzles, Ancient & Modern ", (1997)
"The Magic Garden of George B. and Other Logic Puzzles ", (2007, 2nd ed. 2015)
"King Arthur in Search of His Dog and Other Curious Puzzles ", (2010)

"Alice in Puzzle-Land", (1982a) is *perhaps* the best.

To give you an impression, here's the beginning of the first of the books cited above:

> **1. Was I Fooled?**
> My introduction to logic was at the age of six. It happened this way: On April 1, 1925, I was sick in bed with grippe, or flu, or something. In the morning my brother Emile (ten years my senior) came into my bedroom and said: "Well, Raymond, today is April Fool's Day, and I will fool you as you have never been fooled before!" I waited all day long for him to fool me, but he didn't. Late that night, my mother asked me, "Why don't you go to sleep?" I replied, "I'm waiting for Emile to fool me." My mother turned to Emile and said, "Emile, will you please fool the child!" Emile then turned to me, and the following dialogue ensued:
>
> Emile: So, you expected me to fool you, didn't you?
> Raymond: Yes.
> Emile: But I didn't, did I?
> Raymond: No.
> Emile: But you expected me to, didn't you?
> Raymond: Yes.
> Emile: So I fooled you, didn't I!
>
> Well, I recall lying in bed long after the lights were turned out wondering whether or not I had really been fooled. On the one hand, if I wasn't fooled, then I did not get what I expected, hence I was fooled. (This was Emile's argument.) But with equal

> reason it can be said that if I was fooled, then I *did* get what I expected, so then, in what sense was I fooled. So, was I fooled or wasn't I? (1978, 3–4)

5.6 The Sorites paradox

Here's a way of showing that everyone is bald, no matter how hairy their head is:

1 A person with 0 hair is bald.
2 If a person with n hairs is bald, then a person with $n+1$ hairs is bald too.
3 A person with a billion hairs is bald.

What makes this argument really problematic is the possibility of arguing in the reverse direction thus:

1 A person with a billion hairs is not bald.
2 If a person with n hairs is not bald, then a person with $n-1$ hairs is not bald either.
3 A person with a billion hairs is bald.

Both arguments are inductive: we have an anchor for the induction (**1**), which is as uncontroversial as it gets, and inductive step **2**, the plausibility of which relies, apart from the supposition (which we may make in this context) that nothing else matters to baldness than the number of hairs, on the crucial assumption that small differences in the number of hairs do not make a difference with respect to baldness. This crucial premise is sometimes taken to be the hallmark of predicates that are vague.

$$F \text{ is vague} \quad :\Leftrightarrow \quad \exists x(\neg\mathbf{Det}(Fx) \wedge \neg\mathbf{Det}\neg(Fx))$$

"**Det**" here stands for the "it is determinately the case …" operator, and the x in question is called a "borderline case" for F. The intuitive idea is that a sharp predicate G divides its range of applicability into two jointly exhaustive groups – those that are G and those that are not G. A vague predicate, by contrast, leaves some cases undecided; it has borderline cases that are neither clearly in one nor clearly in the other group.

There is a live question where the source of vagueness, for some specific group of predicates, lies. Roughly, we can distinguish three types of diagnoses:

- According to the semantic conception of vagueness, vagueness is primarily a linguistic phenomenon: for practical or other reasons, the linguistic community has not specified the meaning, i.e. the range of applicability, of vague predicates fine-grainedly enough: it has left the borderline cases undecided.
- According to the epistemic conception of vagueness, vagueness is primarily an epistemic phenomenon: even though vague predicates do trace sharp boundaries, we are, de facto or for principled reasons, ignorant of these boundaries.
- A third, rather heterodox conception locates vagueness in the world: there are no sharp boundaries for vague predicates to latch on, the world itself leaves some matters of baldness undecided.

The most pressing problem for the semantic conception is to provide a semantics for the undecided cases. Introducing a third truth-value ("undecided", "unsettled") runs the risk of falling prey to higher-order vagueness: not only does the vague predicate itself not trace sharp boundaries (and so leave room for borderline cases), the borderline region itself has vague boundaries (this may be put as the claim that not only F, but also $\mathbf{Det}F$ is a vague predicate). The most popular solution to this problem is to redefine truth rather than to introduce a third truth-value. In such so-called "super-valuationist" theories, truth is a matter not of valuation (attribution of the predicate itself to some individual), but of a super-valuation, attributions of precisified predicates to individuals.

For baldness, e.g., precisified predicates would specify the exact number of hairs. We then call a predicate F super-true of some individual a if all its precisifications are true (in the ordinary sense) of a, super-false if all precisifications are false. The theory may thus preserve bivalence for the precisifications and hence make it super-true (true on all precisifications) that every individual is either F or not F but not both.

On such a semantic theory of vagueness, the truth-conditions imposed on the world by the predication of a vague predicate are only fulfilled if the resulting sentence is super-true and they are only failed if the sentence is super-false. Because the inductive premise of the Sorites argument is super-false, the paradox is blocked. This leaves the diagnostic job, however: why is it that for vague predicates, super-truth and not truth is what's needed? What is the status of such 'precisifications'; are they part of (or at least determined by) the meaning of the vague predicate? Will not the range of acceptable precisifications of a given usage of a vague predicate itself be vague and hence in need of further precisification to avoid higher-order vagueness? Another problem pertains to the status of super-truth: can it really be a truth-predicate if it is not disquotational, i.e. does not license the use of the schema (T) discussed below: even though I can call the sentence "there is an exact number of hairs n such that anyone with n hairs is not bald and anyone with $n-1$ hairs is bald" super-true, I cannot therefrom conclude that there really is such an exact number. Rather, "[p]erhaps the ordinary concept of truth *should* match the vagueness of the sentences to which it is applied." (Williamson, 1995, 163)

The main advantage of the epistemic conception is that it straightforwardly 'solves' the paradox by rejecting the induction premise, while still explaining its attractiveness. The prize to pay is its intuitive implausibility and its rejection of the so-called **KK** principle, that everything known is known to be known.

6 Paradoxes of Rationality

6.1 Surprise Exam

A reliable teacher announces there will be a surprise exam on one weekday of the following week. The pupils reason that it can't be on Friday, since if it hasn't come by Thursday evening they will expect if the following day, and then it won't be unexpected. If it hasn't come by Wednesday evening, they will rule out Friday for the reason just given: but then it won't be a surprise on Thursday and so that day is ruled out too. And so on backwards through the week. So the teacher's announcement cannot be fulfilled.
But surely there can be a surprise exam.

The pupils reason inductively, but for their 'backwards' induction to get started they must be sure that there will be an exam. If they doubt that there will be an exam during that week, they cannot be sure that the teacher would have to hold it on Friday if he has not done it on Thursday.

The crucial step in their reasoning is that "if there cannot be a surprise exam on Friday, then there cannot be a surprise exam on Thursday". This induction step, however, must not just be true but also be known by the pupils to be true. It is thus required, for the reasoning to get started, that the pupils know on Wednesday evening that the teacher will keep one's word and really give them an exam and that he will not give it to on a day they expect it to be held. For if they suspect that the exam could be held but not be a surprise, they cannot rule out Thursday.

One possible solution has been to deny the pupils' this piece of knowledge. As the week goes on, they become unable to trust their teacher, because the teacher's statement become more and more alike to the Moore-paradoxical (cf. below) statement "there will be a test tomorrow and you do not expect this".

6.2 Prisoners' Dilemma

We have been arrested for a serious offence and put in separate cells, and we both know that our fate will be determined in the following way:

- If one confesses while the other stays silent, the confessor goes free and the other is sentenced to ten years.
- If both confess we both get seven years. If both stay silent, we both get a year for a lesser offence.

Assume that we both want to minimize our sentences and that we are both rational. Then I will confess. For, if you confess then I had better do so since otherwise I get ten years, and if you stay silent I will go free if I confess. So whatever you do I am better off confessing. Since, given the assumption of rationality, you will decide in the same way, we will both confess and get seven years each. Yet we would both have been better off if we had stayed silent and only got a year each.

The paradox has been said to illustrate the difficulty of achieving coordination:

> Under the assumptions, it is impossible to achieve the full benefits of cooperation unless both parties take a foolish risk. (Clark, 2002, 151)

Many have tried to improve on this pessimistic conclusion by tweaking the initial situation: in an iterated prisoners' dilemma, it has been said, it may become rational to take the "foolish risk" if this may signal the other a general willingness to cooperate. Similarly, Pettit (1986, 182) has argued that it is rational to *hope* that the rational thing to do in the prisoners' dilemma were to cooperate. He gives the following argument:

- If we each act rationally, and our situations are symmetrical, then if I cooperate it is rational for me to cooperate and if it is rational for me to cooperate then the other cooperates too.
- If we each act rationally, and our situations are symmetrical, then if I defect it is rational for me to cooperate and if it is rational for me to defect then the other defects too.
- But we do each act rationally, and our situations are symmetrical.
- Therefore if I cooperate it is rational for me to cooperate and if it is rational for me to cooperate then the other cooperates too; and likewise if I defect it is rational for me to defect and if it is rational for me then the other defects too.
- Hence, I get the better result if I cooperate, assuming it is rational to cooperate; so, being rational, I should hope that it is.

6.3 Newcomb's Problem

It has been argued that the prisoners' dilemma is just a special case of a more intractable problem, which serves to illustrate an important clash of theories in decision theory.

Before you are two boxes: a transparent one containing 10,000 JPY and an opaque one which contains 1,000,000 JPY or nothing. You have a choice between taking the opaque box alone or taking both of them. A Predictor with a highly successful record predicted whether you are going to take both boxes or just one. If he predicted that you will take just the opaque box he has put a million in it: if he predicted you will take both boxes he has left the opaque box empty. And you know this. Should you take one box or two?

We reach a paradoxical conclusion by the joint acceptance of two contrary arguments which seem independently plausible:

- Evidentialist decision theory: you should limit yourself to the opaque box, because generally you should choose the option which has the highest *expected value*. This value of the one-box option is 1 mio because you have reason to assume that, if you one-box, the predictor has predicted this and put 1 mio into the opaque box. In contrast, by parallel reasoning, the expected value of two-boxing is a mere 10'000.
- Causal decision theory: whatever the predictor predicted, the money is now in the opaque box or it is not; in both options, you take the opaque box; the only difference concerns whether or not you also take the transparent box which you know to contain 10'000 – it would be foolish not to take them, taking them is the *dominant strategy*.

As the prisoners' dilemma, Newcomb's problem has many analogies and applications. I find it troubling that different analogies pull me in different directions (cf. (cf. Clark, 2002, 128)).

For an analogy in support of two-boxing, suppose that a certain career-choice is correlated, perhaps even strongly correlated with a certain genetic disease. You do not know whether you have the disease and you hope you don't. Does this give you a reason not to take up the career in question? Surely not!

For an analogy in support of one-boxing, suppose that expansion of the money supply brings down unemployment (good) but only if it is not expected by the markets; if it is, it rather causes inflation (bad); if it falsely expected, the result is even worse (recession).

6.4 Sleeping Beauty

On Sunday Beauty learns she is going to be put to sleep for the next two days, and be woken briefly either once or twice. A fair coin is tossed: if it lands Heads she will be woken just on Monday, if Tails on both Monday and Tuesday. If she is woken just on Monday, she is put back to sleep with a drug that makes her forget that waking. Beauty knows all this, so that when she wakes on Monday she doesn't know which day it is. On waking, what probability should she assign to the coin having landed Heads?

- A half, because it was a fair coin and she's learned nothing new relevant to how it fell.
- One-third, because if the trial were repeated week after week, she should expect twice as many Tails-wakings as Heads-wakings. For every time the coin lands Tails she is woken twice, as compared with once when it lands Heads.

7 Paradoxes of the Mind

7.1 The Paradox of Analysis

For a long time, a dominant view about philosophy was that philosophy deals (mostly, perhaps even exclusively) in conceptual analyses. Such an analysis is done in the armchair and is meant to 'reduce' a philosophically difficult and interesting concept into its conceptual 'components', like the (itself mythological) 'justified true belief' account of knowledge:

(1) x knows that $p :\iff p \wedge x$ believes that $p \wedge x$'s belief that p is justified

The paradox of analysis is that it is hard to see how such an analysis may be true and informative.

- If it true, then it is analytically or conceptually true, but then (1) should be apparent / obvious / known by anyone competent with the concept expressed by "knows".
- If it is informative, it should be surprising / an open question / an interesting thesis that the biconditional expressed by (1) is true; but then the usual philosophical ('armchair') arguments in favour of (1) do not suffice to establish its truth.

Clark (2007) puts the matter in terms of a successful analysis being a statement of (conceptual?) identity:

> We can analyse the notion of brother by saying that to be a brother is to be a male sibling. However, if this is correct, then it seems that it is the same statement as 'To be a brother is to be a brother'. Yet this would mean the analysis is trivial. But surely informative analysis is possible?

7.2 Berry's Paradox

The least integer not describable in fewer than twenty syllables' is itself a description of nineteen syllables. So the least integer not describable in fewer than twenty syllables is describable in fewer than twenty syllables, because the quoted expression is such a description and has only nineteen syllables.

7.3 Moore's Paradox

There are many sentences which are true but which I do not believe. I am not omniscient after all. Let "p" be some such sentence, true, but not believed. Suppose I make the following assertion:

(2) p but I do not believe that p

Moore's "paradox" (or puzzle) is that (2) seems unassertable. It stands at the centre of the discussion about so-called "norms of assertion", e.g. the question whether you should assert only what you know, what you believe or what you (believe to) have reasons to believe.

7.4 The Paradox of Knowability

The so-called "Knower" is a Liar-type sentence about knowledge:

(K) I know that (K) is false

Do I know (K)? If I do, then K is true (I only know true things), hence I know that K is false. But then K is false and I do not know it (for I can know only true things).

If (K) is true, then it is false, and known to be false. If (K) is false, on the other hand, I do not know that it is false. But we have just showed that I *do* know that it is false.

The problem with (K) is not just, as with the Liar, that it is both true and false and cannot be both. It's that it is both known and not known, a contradiction not just for the (supposed) knower but for us who describe her:

- Suppose (K) is true. Then I know that (K) is false, by definition of (K). Then it is false, because knowledge entails truth. Hence, conditionalising: If (K) is true it is false. Hence (K) is false. Hence I do not know that (K) is false, because this is what (K) says.

- By assuming it's truth, I have shown that (K) is false. Hence I have shown that (K) is false. Hence I do know that it is false, for what I have proven I know.

We can construct a contingent version of the Knower paradox as follows. Let there be two teachers, X and Y, making the following statements:

- X: "What Y will say next you can know to be false."
- Y: "What X just said is true."

Both statements, on their own, are intelligible and may well be true. Together, however, they are paradoxical: if what Y said is true, then what X said is true and hence what Y said is false. It follows, by *reductio*, that what Y said is false. What we have just found out is something we know: we know that what Y said is false. So what X said is true. So we can know that what Y said is false. So it is not the case that what X said is true.

It appears doubtful that this is just the problem of the Liar. To show this, Sainsbury (1995, 101) compares (K) to the "Believer":

(B) Mariko does not believe what (B) says.

Does Mariko believe (B)? If she does, she can see that she is believing something false. But then she does not believe it. If she does not believe it, (B) is true and she can see that it is. So she does believe it.

7.5 The Lottery

Suppose there is a lottery with a million tickets. If you bought one, it is reasonable to assume that it will not win. But that is true of all ticket buyers: all of them have reason to assume that they won't win the lottery. But someone will, and we know this.

7.6 The Preface

It is perfectly customary to apologise in the preface of a book for any errors the book might contain and even to say that there are some such errors. Once such a humble acknowledgment of fallibility is included in the book, however, it can no longer be jointly true: at least one affirmation in the book must be false.

7.7 The Good Samaritan

If you commit murder you ought to do so gently. Suppose then that you do commit murder. Then you ought to do so gently. Now from you ought to do A it follows that you ought to do anything logically implied by A. (For example, from I ought to help my mother and spend time with my children it follows that I ought to help my mother.) You are murdering gently entails you are committing murder. So it follows that you ought to commit murder.

One important distinction here is between 'broad-' and 'narrow-scope' obligations. The distinction may be exemplified, in the epistemic case, by two readings of:

(3) I thought your yacht was bigger than it is.

On one – 'broad'-scope – reading, (3) says I believed something logically impossible, namely that the size of your yacht is different from itself. The more plausible 'narrow-scope' reading says that there is a size, the one your yacht actually has, which I believed not to be the size of your yacht.

7.8 The Toxin Paradox

You are offered a million dollars to form the intention of drinking a vile potion which, though not lethal, will make you unpleasantly ill. Once you have formed the intention the money is handed over, and you are free to change your mind. The trouble is that you know this, and it will prevent you from forming the intention, since you cannot intend to do what you know you will not do.

The problem is, as Clark (2007, 195) says, that " "I intend to drink the potion but I won't drink it" is as self-defeating a belief or utterance as one of the form "p but I don't believe it"'.

7.9 The Eclipse Paradox

During a total solar eclipse we see the moon as a dark round disc silhouetted against the sun. But which side of the moon do we see? Standardly, we do not see an object unless the object causes our perception. But in this case it is the far side of the moon which absorbs and blocks the sunlight from us and causes us to see the moon as a dark disc. So during an eclipse it is the far side of the moon, not its near side, that we see. But that seems quite contrary to the way we think of seeing.

8 Problems

Problem 1. Let us suppose that Akitans always tell the truth and Tokyoites always lie and that all people mentioned in this example are from either one of these two cities.

(a) We meet a person A who says of herself and of another person B: "Either I am from Tokyo or B is from Akita." Wherefrom are A and B?

(b) We meet three people, A, B and C, two of whom make the following assertions:

A: "All the three of us are from Tokyo.".

B: "There is exactly one Akitan among us.".

Where are they from?

(c) Another trio, A, B and C. A says that B and C come from the same city. We then ask C: "Do A and B come from the same city?" What will she answer?

(d) Suppose that A says: "Either I am from Tokyo or 2+2=5." What can we deduce from this?

Problem 2. A logic student meets someone very interesting. She asks him: "Will you do me a favour? I will make a statement. All I ask you is that if this statement is true, you will give me your phone number." – He agrees. – "And there is another favour I wanted to ask you: if my statement is false, you will not give me your phone number. " – And he agrees to this too.
The logic student then makes her statement and her interlocutor discovers, much to his surprise, that to keep his promise, he must give her a kiss!
What could be the statement that has such magical powers?

Problem 3. Let us suppose that Akitans always tell the truth and Tokyoites always lie and that all members of a family come from the same city.

(a) Two brothers from one of the two cities were asked whether they were married. Their answers were:

Kentaro "We are both married."

Ryo "I am not married."

Is Kentaro married? What about Ryo?

(b) The same question was asked to two sisters, who answered:

Akiko "Either both of us are married or none is."

Misako "I am not married."

Who is married?

(c) The same question was asked to a third duo. The older of two brothers has only said that at least one of them was married. Unfortunately, all records of what the younger has said were lost, but we know at least that either he said that he was married or that he said that he was not married. We do know, however, that based on this information, their interlocutor was able to figure out who of the two was married. Do you know which one?

Problem 4. Let us suppose that Akitans always tell the truth and Tokyoites always lie.

(a) What assertion can be made by an Akitan woman or an Tokyoite man, but cannot be made by a Akitan man or a Tokyoite woman?

(b) What assertion can be made by any woman, Akitan or Tokyoite, but cannot be made by any man, Akitan or Tokyoite?

(c) What assertion can only be made by a Akitan woman, but not by an Akitan man nor by any Tokyoite?

(d) What assertion can only be made by a Tokyoite woman?

Problem 5. Ski camp: there is a nocturnal control of sleeping arrangements in the ski camp. Three sleeping bags are found, with two people inside each of them. The instructors know that in one of them, there are two men, in another one there are two women and in a third one, there are a man and a woman. They randomly choose one sleeping bag, open it, and find one man (the other person still being hidden inside). What is the probability that the other person in that bag is also a man?
The chief instructor, Haiko, reasons as follows:

> The probability that the other person in this bag is a man is 50%. Because we have already found a man inside, it cannot be the all-female sleeping bag. So it must be one of the two others. In these two other cases, either a man or a woman is the second person.

Is he right?

Problem 6. There are three boxes, one with a prize in it, the other two empty. Here is a piece of reasoning that seems to show that the probability of finding the prize in a given box is 1/2:

Let us call the three boxes A, B and C. Suppose we have chosen A and think about the (certain!) fact that one of the other two boxes is empty:

P1 If B is empty, then the probability that the prize is in A is 1/2.
(Because it's either in A or in C.)

P2 If C is empty, then the probability that the prize is in A is 1/2.
(Because it's either in A or in B.)

We note that the disjunction is exhaustive: either B or C must be empty! So we have:

P3 Either B is empty or C is empty.

By eliminating the disjunction, we conclude that the probability that the prize is in A is 1/2.

Where lies the fault?

9 Solutions

Problem 1. The Tokyoites who always lie and the Akitans who always tell the truth (after Smullyan (1978, 21–25)).

(a) If A is a Tokyoite, then A has lied: then *neither* A is a Tokyoite *nor* is B Akitan. Under this supposition, in particular, A is both Tokyoite and not Tokyoite. Hence A is not Tokyoite and must be Akitan. Hence he has told the truth and B is Akitan too.

(b) A cannot be right: if he were Akitan, he would have told the truth, so he would be Tokyoite. Hence A is Tokyoite and there must be at least one Akitan among them. If B is also a Tokyoite, then C must be Akitan, but then what B says is true and B is *not* a Tokyoite. Hence B must be Akitan and tells the truth. Hence C must be a Tokyoite.

(c) If A is Akitan, he tells the truth and B and C are from the same city. If C is Tokyoite (and B is too), C will answer "yes", because A and B are not from the same city and C is always lying. If C is Akitan (and B is too), he will answer "yes", because A and B are of the same city and C never lies.
If, on the other hand, A is Tokyoite, A has lied and B and C are not from the same city. If C is Tokyoite (and hence B Akitan), C will answer "yes", because A and B are not from the same city and C is always lying. If C is Akitan (and hence B Tokyoite), then C will answer "yes" because A and B are from the same city and C never lies. In all cases, C will answer "yes".
In all four possible cases, therefore, C will answer "yes".

(d) The only possible conclusion is that the supposition is false: A cannot have said this. If he had said it and were Akitan, then $2 + 2 = 5$, which is impossible. If he had said it and were Tokyoite, he would not have lied, which is also impossible.

Problem 2. How logic helps to flirt (from Smullyan (1997)). A sentence that will do the job is (there are many others):

(D) You will neither give me your number nor a kiss.

If (**D**) is true, then he will have to give his phone number (that was the first promise). But then (**D**) would be false. Hence (**D**) must be false, i.e. the logic student will either get his phone number or a kiss. But he cannot give her his number (that was the second promise). So he must kiss her to keep his word.

Problem 3. Lying Tokyoites and honest Akitans again.

(a) Because Kentaro and Ryo are brothers, either they both tell the truth or they both lie. Because the two statements are in conflict with each other, they cannot both be true. Hence the brothers are both lying: Ryo is married but Kentaro is not.

(b) If both statements are true, both sisters are unmarried. If both statements are false, then Misako is married but Akiko is not. In both cases, Akiko is unmarried, but we cannot say anything about the marital status of Misako.

(c) If the younger brother said that he was married, then it is both possible that both statements are false (and hence that both brothers are unmarried) and it is possible that both statements are true (and hence the younger brother married and the older either married or unmarried). So the interlocutor would not be able to draw any conclusion.

If, on the other hand, the younger brother said that he was unmarried, it is not possible that both assertions are false: if the younger brother's statement is false, then the older brother's statement must be true. Hence in this second case, both statements must be true, which means that the younger brother is unmarried and the older brother married.

Problem 4. Here are some sentences that do the job (there are many others):

(a) "I am a woman" – made by an Akitan woman, the assertion is true; made by a Tokyoite man, it is false. The assertion cannot be made by an Akitan man (because it would be false) nor by a Tokyoite woman (because it would be true).

(b) "I am either an Akitan woman or a Tokyoite man" – made by an Akitan woman, the assertion is true; made by a Tokyoite woman, the assertion is false. The assertion cannot be made by any man from any of the cities.

(c) "I am not an Akitan man" – true if made by an Akitan woman, false if made by an Akitan man and true (and hence impossible to make) for any Tokyoite.

(d) "I am a Tokyoite man" – the Akitans cannot make this assertion (because it would be false), and Tokyoite men cannot make it because it would not be false.

Problem 5. Haiko makes a mistake which is quite common in probabilistic reasoning. In fact, the probability is 2/3, i.e. 66.67 %. We can argue for this the following way:

1. Forget about the all-women sleeping bag. In the two others, there are three men, let us call them A, B and C. Suppose that A and B share a sleeping bag and that C is with a woman. The man we have found may, with equal probability, be any one of A, B or C. If he is A or B (which is the case with a probability of 2/3), then the other person in the sleeping bag is also male.
2. If we forget about the all-women sleeping bag, four persons are involved, three men and one woman. One of the men we have found. The unknown person sharing a sleeping bag with him may be any of the other three, two of which are men, which gives a probability of 2/3.
3. Let us change the example and take four cards, three red ones and one black one. If I pick one randomly and its red, there are three cards left, two red and one black. If I now pick one of them at random, the probability of my picking a red one is *obviously* 2/3.

Haiko makes the mistake of being misled by an apparently binary opposition (male vs. female), whereas really there are three options to consider.

Problem 6. Reasoning about the boxes:

P1 If B is empty, then the probability that the prize is in A is 1/2.
(Because it's either in A or in C.)

P2 If C is empty, then the probability that the prize is in A is 1/2.
(Because it's either in A or in B.)

P3 Either B is empty or C is empty.

P3 The probability that the prize is in A is 1/2.

Their appearance notwithstanding, **P1** and **P2** are not best interpreted as material conditionals. Rather, they are statements of conditional probability:

P1′ In half of the cases where B is empty, the prize is in A. To have six equiprobable cases,

P2′ In half of the cases where C is empty, the prize is in A.

we need to add:

P4 In both halves of the cases where A is empty, the prize is not in A.

C′ In two of six equiprobable cases, the prize is in A (the probability is 1/3).

The mistake may be due to the fact that the probability 'operator', which is in effect conditional, i.e. talks only about proportions among cases where the antecedent is true, is misleadingly placed into the consequent. Less misleadingly, the premisses are put thus: **P1″** The probability is 1/2 that: if B is empty,

P2″ The probability is 1/2 that: if C is empty,

This is a familiar phenomenon in natural language:

conditional bets "If you're playing another game, I bet you'll win" does not mean that I only make a bet if you're playing, it means that I bet that: if you play you win.

conditional promises "If you trust me I promise that I won't let you down" makes (unconditionally) a promise, even if you do not trust me. The assertion gives you the choice between lack of trust and expectation of non-deception.

conditional insults "If he does not prepare, the idiot will fail the exam" is an (unconditional) insult. Whether or not he prepares, the poor guy is called an idiot for not succeeding in the exam unless he prepares (one is led to wonder how else he should be succeeding).

References

Black, Max (1951). Achilles and the Tortoise. *Analysis*, 11(5): 91–101. Reprinted in Black (1954, 95–108).

Black, Max (1954). *Problems of Analysis: Philosophical Essays*. Ithaca, New York: Cornell University Press.

Cantor, Georg (1892). Über eine elementare Frage der Mannigfaltigkeitslehre. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 1: 75–78. Translation by William Ewald: 'On an elementary question in the theory of manifolds' in Ewald (1996, 921–922).

Clark, Michael J. (2002). *Paradoxes from A to Z*. London: Routledge. Second edition: Clark (2007).

Clark, Michael J. (2007). *Paradoxes from A to Z*. London: Routledge, 2nd edition. First edition: Clark (2002).

Ewald, William Bragg, ed. (1996). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, *Oxford Science Publications*, volume II. Oxford: Oxford University Press.

Frege, Gottlob (1884). *Die Grundlagen der Arithmetik: eine logisch-mathematische Untersuchung über den Begriff der Zahl*. Breslau: Wilhelm Koebner. Reissued as Frege (1961).

Frege, Gottlob (1893). *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*, volume I. Jena: Hermann Pohle. Reissued as Frege (1966).

Frege, Gottlob (1961). *Die Grundlagen der Arithmetik*. Hildesheim: Georg Olms Verlagsbuchhandlung.

Frege, Gottlob (1966). *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*, volume I. Hildesheim: Georg Olms Verlagsbuchhandlung. Reprografischer Nachdruck der Ausgabe Frege (1893).

Haslanger, Sally (2003). Persistence Through Time. In Loux, Michael J. & Dean W. Zimmerman, eds., *The Oxford Handbook of Metaphysics*, Oxford Handbooks, pp. 315–354. Oxford: Oxford University Press.

van Heijenoort, Jan, ed. (1967). *From Frege to Gödel: A Source Book in Mathematical Logic 1879-1931*. Cambridge, Massachusetts: Harvard University Press.

Horwich, Paul (2016). The Nature of Paradox. In Ostertag, Gary, ed., *Meanings and Other Things. Themes from the Work of Stephen Schiffer*, pp. 211–228. Oxford: Oxford University Press.

Lycan, William G. (2010). What, exactly, is a paradox? *Analysis*, 70(4): 615–622.

Peirce, Charles Sanders (1935). *Collected Papers of Charles Sanders Peirce: Scientific Metaphyiscs*, volume 6. Cambridge, Massachusetts: Harvard University Press. Edited by Charles Hartshorne and Paul Weiss.

Pettit, Philip (1986). Preserving the Prisoner's Dilemma. *Synthèse*, 68: 181–184.

Quine, Willard van Orman (1962). Paradox. *Scientific American*, 206(4): 84–99. Reprinted as "The Ways of Paradox" in Quine (1966, 1–18).

Quine, Willard van Orman (1966). *The Ways of Paradox and other essays*. Cambridge, Massachusetts: Harvard University Press. Second, enlarged, edition: Quine (1976).

Quine, Willard van Orman (1976). *The Ways of Paradox and other essays*. Cambridge, Massachusetts: Harvard University Press, 2nd edition. First edition: Quine (1966).

Russell, Bertrand Arthur William (1903). *The Principles of Mathematics*. Cambridge: Cambridge University Press. 2nd edition: Russell (1937).

Russell, Bertrand Arthur William (1910). *Philosophical Essays*. London: Longmans, Green & Co. Second edition: Russell (1917).

Russell, Bertrand Arthur William (1914). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. LaSalle, Illinois: Open Court Publishing Co. Revised edition: Russell (1926).

Russell, Bertrand Arthur William (1917). *Mysticism and Logic*. London: George Allen & Unwin. Second edition of Russell (1910).

Russell, Bertrand Arthur William (1926). *Our Knowledge of the External World as a Field for Scientific Method in Philosophy*. London: George Allen & Unwin, 2nd edition. Revised edition of Russell (1910).

Russell, Bertrand Arthur William (1937). *The Principles of Mathematics*. London: George Allen & Unwin, 2nd edition.

Sainsbury, Richard Mark (1995). *Paradoxes*. Cambridge: Cambridge University Press, 2nd edition.

Salmon, Wesley C. (2001a). Introduction. In Salmon (2001b), pp. 5–44.

Salmon, Wesley C., ed. (2001b). *Zeno's Paradoxes*. Indianapolis, Indiana: Hackett Publishing Co.

Schiffer, Stephen (2003). *The Things We Mean*. Oxford: Oxford University Press.

Smullyan, Raymond M. (1978). *What is the Name of this Book? The Riddle of Dracula and Other Logical Puzzles*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Smullyan, Raymond M. (1980). *This Book Needs No Title. A Budget of Living Paradoxes*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Smullyan, Raymond M. (1982a). *Alice in Puzzle-Land. A Carrollian Tale for Children Under Eighty*. London: Penguin Books.

Smullyan, Raymond M. (1982b). *Lady or the Tiger? And Other Logic Puzzles Including a Mathematical Novel That Features Gödel's Great Discovery*. New York: Times Books.

Smullyan, Raymond M. (1983). *5000 B.C. and other Philosophical Fantasies*. New York: St. Martin's Press.

Smullyan, Raymond M. (1992). *Satan, Cantor, and Infinity and Other Mind-Boggling Puzzles*. New York: Alfred A. Knopf.

Smullyan, Raymond M. (1997). *The Riddle of Scheherazade: And Other Amazing Puzzles, Ancient & Modern*. San Diego, California: Harvest.

Smullyan, Raymond M. (2007). *The Magic Garden of George B. and Other Logic Puzzles*. Monza: Polimetrica. Reissued as Smullyan (2015).

Smullyan, Raymond M. (2010). *King Arthur in Search of His Dog and Other Curious Puzzles*. Mineola, New York: Dover Publications.

Smullyan, Raymond M. (2015). *The Magic Garden of George B. and Other Logic Puzzles*. Singapore: World Scientific Publishing Co. Reissue of Smullyan (2007).

Thomson, James F. (1954). Tasks and Supertasks. *Analysis*, 15: 1–13.

Whitehead, Alfred North (1929). *Process and Reality. An Essay in Cosmology*. London: MacMillan Publishing Co. Gifford Lectures Delivered in the University of Edinburgh During the Session 1927-28; corrected edition: Whitehead (1978).

Whitehead, Alfred North (1978). *Process and Reality. Corrected Edition*. London: MacMillan Publishing Co. Edited by David Ray Griffin and Donald W. Sherburne; first edition: Whitehead (1929).

Williamson, Timothy (1995). Is Knowing a State of Mind? *Mind*, 104: 533–563.