Week Twelve: Ramsey Sentences and Moral Realism

Apology: By a terrible misfortune, parts of the handout went missing *en route* from my home computer to my work computer. (I.e. I transferred the wrong file ☹☹☹.) So the last couple of sections are a bit more point-form-ish than we may have wanted. I hope this won't be too much of a distraction.

Chapter 5

1. Descriptive and ethical properties

Chapter 5 of **Jackson** is largely about the relationship between descriptive and normative properties. Jackson first says that the ethical properties supervene on the descriptive properties, and then notes there is a two-line proof from this to the entailment of the ethical facts about a world from the descriptive facts. This is relatively easy to follow, but when he tries to parlay this into an argument for the *identity* of certain ethical properties with certain descriptive properties. So we'll start with some effort to be clear about this.

1.1. Identity Claims

"I will start by arguing that the nature of the supervienience of the ethical on the descriptive tells us that ethical properties are descriptive properties in the sense of properties ascribed by language that falls on the descriptive side of the famous is–ought divide." (118-9)

It is hard to know quite what to make of this identity claim. Here's a bad conclusion to draw from it. Some argue that we cannot give an analysis of ethical concepts such as being good in terms of descriptive concepts. Well, if Jackson is right then there is such an analysis: *x* is good iff *x* is good. The right-hand side is descriptive, so we have an analysis of being good in terms of descriptive concepts. Something has gone seriously wrong here, I suppose. (Maybe this is a false assumption).

Also note quickly that the line between the ethical and the descriptive is notoriously hard to draw in practice. Consider, to just give one example, the property of being alive. There are notorious hard cases about where the border between being alive and not being alive is at either end of one's life. And clearly there are ethical implications which arise from particular drawings of the line. Whatever moral duties we have to those things which might become alive, or once were alive, they pale in comparison to the duties we have to the living. But it isn't clear, well not to me at least, whether that is because there are certain bridge principles from facts about life to ethical duties, or because 'alive' is itself a part of *ethical* language. (Remember the difficulties we had last week in trying to precisely define 'life'.) I don't even know quite what is at issue here, so we shouldn't be too flippant about assuming there is a clear distinction here. I will try to be just a little bit flippant.

1.2.   The Argument for the Identity Theory

Jackson doesn't directly argue for the identity of ethical and descriptive predicates. First, he argues that all sentences using ethical terms are metaphysically equivalent to some sentences involving only descriptive language. The form of argument Jackson uses here should be familiar from earlier chapters; it is just a minor variant on the argument he uses for the 'entry by entailment' thesis in chapter one. (This argument appears on pages 122-3)

We have already assumed that ethical sentences are true or false, for short, are truth-apt. So most of them will be true at some worlds. The one's that are not are equivalent to the descriptive sentence: "There are electrons or there are no electrons." Otherwise, assume we have some sentence E framed partially in ethical terms. Then for every world w_S in which E is true, there is a maximal descriptive 'sentence' D_S which is true in w_S. Jackson claims that E is equivalent to the disjunction of the various D_S, which we'll call D. It should be clear enough that E entails D; whenever E is true one of the disjuncts of D is true, by definition, and presumably whenever a disjunct is true the disjunction is true. The proof that D entails E relies on the supervenience claim; if D doesn't entail E then it is possible that D is true in some world and E false. But then there is some D_S which is true in one world where E is true (or else it would not be a disjunct) and true in another world where E is false (or else it would not show that this possibility holds). Given the supervenience claim this is impossible, so D and E are logically equivalent.

Jackson then says that "the same line of argument can be applied *mutatis mutandis* to ethical and descriptive predicates and open sentences: for any ethical predicate there is a purely descriptive one that is necessarily co-extensive with it." Actually the argument is a bit more slippery, which is presumably why Jackson has not given it. The problem is that even though there is a reasonably widespread agreement that the totality of ethical facts supervene on the totality of descriptive facts, there is much less widespread agreement on the partial supervenience claims. For instance, do the ethical facts about this room supervene on the descriptive facts about this room. Probably not, there are various ways the world could be that would compel us to take a more active role in the world than others. Whether sitting around discussing supervenience claims is a morally permissible way to pass the time, as opposed to just being pleasant, depends on how much good we could do, or evil we could prevent, outside here. (Maybe not; maybe duties to prevent evil supervene on the evidence we have about the preventable evil, and maybe that does supervene on brain states. But you can see the potential problem.) Now someone trying to run Jackson's argument for properties has a tough choice to make: either come up with a plausible 'partial' supervenience thesis to fill for the global supervenience thesis, or concede that *any* ethical sentence is really making a claim about all of the universe. I don't know which way to jump here, or how much it would hurt to jump either way.

Even just concentrating on the specific argument that Jackson gives for sentences, there are a couple of premises which we might question. First, it isn't clear that there will be a descriptive sentence which 'gives the nature [of a world] in full'. Presumably this is just what Jackson's opponents deny. I manage to not believe this some days because I forget what sentences with uncountably many conjuncts mean. But ignoring this techie problem, I don't know what Jackson has to say to the person who simply denies that the descriptive sentences he wants could not exist

because a full description of the world includes its ethical nature and any descriptive sentence must leave this out. So maybe Jackson's argument here is question-begging against such an opponent.

Secondly, and this really is a techie worry, even if we agree that each D_S is a descriptive sentence and a full description of a world, it isn't obvious that D is a descriptive sentence. Remember that D is (a) a very long disjunction of descriptive sentences and (b) is metaphysically equivalent to, say, "Killing sharks for food is morally permissible." Jackson implicitly relies here on a closure principle; and disjunction of descriptive sentences is a descriptive sentence. This seems plausible for finite disjunctions, but when there are beth-several disjuncts, or some other large infinity, I don't feel much intuitive pull behind the principle. And I feel some pull the other direction — it entails that there is a descriptive sentence equivalent to "Killing sharks for food is morally permissible." Again, my deepest worry here is not that Jackson has made an assumption which is false, it is that Jackson has made an assumption whose truth value we have no means for detecting.

1.3. Senses and Identities?

We noted earlier that there was some difficulty with identifying ethical and descriptive properties. In particular, this identification seems to make it too easy to give a descriptive analysis of ethical properties. Maybe the problem can be solved if we say that some ethical properties, while necessarily co-extensive with descriptive properties, nevertheless have different senses. This will be a problem if we identify sense with intension, i.e. reference in all possible worlds, so this solution may be more trouble than it is worth. Anyway, this solution doesn't seem vulnerable to the Occamist objection on page 127; we have a reason why some descriptive properties should have distinct senses when viewed through an 'ethical' framework.

1.4. Asymmetry

Here's a consequence of Jackson's arguments about descriptive and ethical properties. Some descriptive properties are really ethical. Doesn't that look like a set back for physicalism. Maybe we can go further. Maybe there are only ethical properties. That would be an interesting result.

Well, we can't get that out of Jackson's work. The main reason is that while the ethical supervenes on the descriptive, the descriptive does not supervene on the ethical. Two possible worlds which just consist of rocks flying aimlessly about are ethically equivalent, but not physically equivalent.

2. Moral functionalism

We've seen so far that that ethical properties are descriptive properties, provided we give a suitably deflationary reading of that identity. Now we'd like to find out which descriptive properties they are. First we'll talk about the philosophical foundations of our search. Then we'll say a bit more about the details of the search.

The theory Jackson endorses about the meaning of the moral terms is thoroughly holistic. Folk morality consists of a network of inter-related principles, some relating ethical and descriptive properties (causing pain is *ceteris paribus* bad) and more relating various ethical properties (courageous

people more often perform good actions than cowardly people). And possibly some of the principles are meta-linguistic (we refer to the same properties by our ethical language). These principles would be very useful for fending off the ambiguity which usually attends network analyses of moral terms.

So we don't just find the functional role that, say, a descriptive property has to fill in order to be *goodness*. Rather, we find a whole host of roles which a sequence of descriptive properties $<d_1, d_2, \ldots, d_n>$ must play (roughly) in order to be the ethical properties $<$goodness, courage, …, evil$>$. This is important to ward off a pressing problem. If we believe some utilitarian theories, then we will be able to give a simple analysis of the core ethical terms in descriptive language. But we don't want to have our semantic theory for moral language held hostage to a somewhat controversial ethical theory. But if we don't assume utilitarianism, it seems we will only be use the principles relating ethical and descriptive properties in defining the functional role for an ethical property like *goodness*. The worry here is that no ethical property will be uniquely characterised by these ethical-to-descriptive principles, so we won't be able to build an analysis of the 'higher-level' ethical terms out of descriptive terms plus 'lower-level' ethical terms which can be directly defined. So we break out of an apparent circle by defining all the ethical terms at once, or at least trying to.

The functional role which a property must play to be, say, *goodness* is not defined by the current precepts of folk morality, but by those of *idealised* folk morality. It isn't entirely clear how this should be defined: John Hawthorne has suggested that it is what we would agree about after a really long chat. Given the depth of the apparent confusions in folk morality (particularly about acts and omissions) we shouldn't expect this qualification to be dropped any time soon.

2.1.   Ecumenicism

Jackson claims it is a large advantage of moral functionalism that it is ecumenical. Just being moral functionalists doesn't commit us to consequentialism or Kantian duties or to whatever virtue ethicists are saying these days. In fact, moral functionalism can nicely state the differences between these positions; they are having disputes over which principles are at the core of folk morality.

Nor does moral functionalism take sides in the dispute between those who think the 'thin' ethical concepts like goodness, badness and so on are fundamental to ethics and those who think 'thick' ethical concepts like honesty, courage, *etc* are fundamental. In fact, moral functionalism can state the differences between these positions. The 'thin' theorists think that the principles relating the thin ethical concepts to descriptive concepts, and to each other, are sufficient to define those concepts; their 'thick' opponents do not.

I suspect functionalism is even more ecumenical than Jackson would like. On page 135 Jackson notes that we can try taking any of the principles of folk morality as primitive and recovering the other principles from them. (He thinks only consequentialists will succeed in this regard, but that is beside the point.) But it isn't clear why Jackson thinks this project, of isolating part of folk morality and recovering the rest from that part, is essential. It doesn't seem to follow from functionalism that there is such a privileged set.

Ramsey Sentences

3. Detectives

Let's tell a detective story. We have come across a crime scene, and Marlowe has just figured out what happened. "Whoever jimmied the window was left-handed; but Lefty didn't have the muscle in case things got out of hand, so he brought some help. After they broke in Help went into the bedroom to take care of the occupants, but only found Miss Smith; Lefty went into the study and collected the documents, then Lefty and Help had an argument, probably because Lefty didn't want Help to hit Smith so hard, and Help punches Lefty, who started bleeding here, Help has smashed open the front door, taken the car and driven off, and Lefty has has to stagger out to the street, still bleeding on the pavement, and walked away."

      Marlowe hasn't committed himself about who Lefty and Help are. Really he has said little more than if he had said, There were two people x and y, and x is left-handed, but didn't have the muscle in case things got out of hand, so he brought some help. Suppose that substituting Sternwood and Bogart for Lefty and Help makes Marlowe's story true. Then we would be inclined to say that (a) Sternwood is Lefty, (b) Bogart is Help and (c) Marlowe's story really is true. Something is kind of odd here. Marlowe never gave a definition of 'Lefty' or 'Help', and certainly we didn't know what they referred to beforehand, yet we can now judge whether or not someone is Lefty.

      Matters are more complex when we notice that Marlowe's story may not be entirely true. Say that some of the documents were taken by Sternwood, but Bogart took more of them after hitting Sternwood. We might think that Marlowe's story isn't quite true, is false that is, but still Lefty is Sternwood and Help is Bogart. So Marlowe may have managed to define these two new terms Lefty and Help, which do refer to real people, by telling a story which is false.

      Lewis suggests that this is a paradigm for the way theoretical terms get introduced in science[1]. We don't have to explicitly say what an electron, or a cell, or a demand curve is, we can just use theoretical terms and if our theory is true enough these terms will refer. If the 'theory' introducing the terms is off the mark by a long way, the theory will not only be false, but the theory won't refer. So if there were four criminals in our story, and the bloodstains Marlowe noticed were caused by one of the perps cutting himself on the window while coming in, probably his terms 'Lefty' and 'Help' are non-referring. We already noticed that the theory can be a little false while the terms still refer, but not too false it seems. So it presumably is *a priori* and necessary that anyone who is Lefty has some properties in common with the character in Marlowe's story.

      Now what does all this have to do with ethics? Well, we need not dissect a story, like Marlowe's story about the crime scene, or Jevons's story about 'marginal utility' into the new terms and the old terms. We may be better off disecting the story into terms which we understand and

---

[1] "How to Define Theoretical Terms" *Journal of Philosophy* 67 (1970): 427-46; "Psychophysical and Theoretical Identifications" *Australasian Journal of Philosophy* 50 (1972): 249-58.

terms which we do not understand. So perhaps we can dissect the story into the descriptive and the ethical.

If we do this there is some hope of getting the kind of necessary *a priori* connections between the ethical and the descriptive that Jackson is looking for. Remember that if we took the terms 'Lefty' and 'Help' to be introduced by Marlowe's story, it became necessary and *a priori* that anything which is Lefty has certain properties. It is contingent and *a posteriori* that something is Lefty, but not that Lefty has these defining properties.

4. How it works for explicit definitions

We'll quickly run through the formalism behind this approach, partially because it is intrinsically interesting, and partially to reassure ourselves that there aren't any tricks being pulled. As noted above, Marlowe's story seems to be equivalent at some level to a quantified story. So if we replaced Marlowe's eloquent tale with 'R Lefty Help', what he said seems to be true iff this is true:

$$\exists x_1 x_2 \, (R x_1 x_2 \wedge \forall y_1 y_2 \, (R y_1 y_2 \rightarrow x_1 = y_1 \wedge x_2 = y_2))$$

The second clause is because it would be odd if there is another pair, as well as Greenstreet and Bogart, who fill the R role. In that case we might be tempted to say Marlowe's story is false. We'll return to this presently. Anyway, given that is Marlowe's story, we can now define what it is to be Lefty:

$$\text{Lefty} = \iota x_1 \, \exists x_2 \, (R x_1 x_2 \wedge \forall y_1 y_2 \, (R y_1 y_2 \rightarrow x_1 = y_1 \wedge x_2 = y_2))$$

'ι' is my poor excuse for a definite description operator. In English, this says that Lefty is the thing such that there is some x_2 such that <Lefty, x_2> uniquely satisfy the R-role. So it is necessary and *a priori* that Lefty, if he exists, satisfies his half of the R-role.

The story we have told so far works for names. But sometimes terms are introduced which are not names. The trick here is to convert every term-introducing sentence so it reads like a name. So instead of having 'Electrons are negatively charged' in the term-introducing theory, we have 'Things which have the property of *being an electron* are negatively charged'. Provided we are happy quantifying over properties (and why shouldn't we be?) the story can be run as before.

We have to make some amendments to the story in order to accommodate the intuition that in some circumstances Marlowe's terms will refer even though his story is false. As can be seen the story so far does not allow this possibility. One way out would be to separate the story into two parts. For reasons that will soon become clear, call these the analytic part and the contingent part. We Ramsify over the analytic part to get the reference of the terms, and we insert those references into the contingent part to determine whether the story is true.

This looks like an embarrassment. One idea behind the Ramsey sentence method was to avoid tricky questions about how to divide facts about meaning from facts about the world. (We've

already seen that this division can be hard to locate.) It seems unlikely that we will be able to avoid all such questions, but we can make some headway.

One approach is to say that the names refer if most of the story is true, or anyway some significantly high percentage. So say that Marlowe just said three things R_1 Lefty Help, R_2 Lefty Help and R_3 Lefty Help, and the names refer if two of these are true. Define a new predicate D such that Dxy is, by definition, $(R_1xy \& R_2xy) \vee (R_1xy \& R_3xy) \vee (R_2xy \& R_3xy)$, and say that Lefty is defined as follows:

$$\text{Lefty} = \iota x_1 \exists x_2 (Dx_1x_2 \wedge \forall y_1y_2 (Dy_1y_2 \rightarrow x_1 = y_1 \wedge x_2 = y_2))$$

So if any pair uniquely satisfy two of the things Marlowe said[2], Lefty is the first member of that pair.

This probably won't do in general. Some of the things are 'more definitive' of Lefty than others. If there is someone who perfectly fills the Lefty role in all respects save that he hailed a cab to get away, rather than walking, that doesn't cast any doubt over the denotation of Marlowe's terms. (I suppose cabbies in Hollywood in the 30's didn't ask embarrassing questions about blood drips, provided you kept the seats clean.) But if we find that the same person pried open the window and attacked Miss Smith, it may be no longer clear what is being referred to.

The best approach seems to be to set some kind of weighting to the parts of the story, and let D be the disjunction of all the parts of the story which are sufficiently weighty. This kind of inegalitarianism may offend some Quineans, but it shouldn't offend Quine. After all, it is just a way of adopting his distinction between more and less central parts of a theory. Or to put it in Stalnaker's slightly more colourful terms, we are claiming there are shade**s** (plural!) of grey, not black meaning postulates and white factual claims. We will get necessary truths as *output*, but not of the kind Quine was suspicious. So we may, as Jackson put it, pay 'due homage' to Quine's critique of analyticity.

5.   Unique Realisations and the Field example

Lewis spends a lot of time in 'Theoretical Terms' arguing that is the term-introducing theory is false if it is multiply realised. So when Jevons introduces the concept of marginal utility, saying things like 'agents buy goods until the marginal utility of buying becomes negative', what he says is false if there are two properties which play the marginal utility role.

Lewis has a positive argument for this and a negative argument to back it up. The positive argument is that reference would have to be arbitrary if multiply realised terms refer. Reference is never arbitrary, so multiply realised terms don't refer. The negative argument is that since we are only requiring unique *actual* realisability, not unique *possible* realisability, this isn't the absurd result that many people take it to be. Neither argument, as it turns out, works.

Katherine Bedard, in a recent paper, argued that we can adopt supervaluational methods to avoid the arbitrariness of multiple reference. So if Peter, Paul and Mary all fill the Lefty role, then sentences about Lefty, say "Lefty went to Harvard Business School" are true if all of them went to

---

[2] There are some scope worries here; I'm sure you can figure them out.

Harvard Business, false if none of them did, and gappy if some did and others didn't. Of course some stories may rule out multiple realisations as part of the story, particularly if Marlowe had said, "Lefty and Help were then alone in the house" and this is a sufficiently weighty part of the story. For the record, Lewis more or less accepts Bedard's move here, except in the case where the realisers are massively different.[3]

The negative argument doesn't fare much better. Field points out that there is no clear best deserver for the Newtonian (or folk) concept of mass in contemporary physics. There are two similar concepts of mass, one defined in terms of total energy, the other defined in terms of non-kinetic energy, which both look like they could play the mass role. But it would be absurd to say that Newton wasn't referring to anything when he talked about 'mass', so unique realisation can't be necessary for successful reference of theoretical terms. Again, Lewis now accepts this argument.

6. How it works for tacit definitions

- The term-introducing story is the list of platitudes in idealised folk morality.
- Not clear how far this will have to be extended. Does it include idealised epistemology?
- This involves taking the myth of our 'Rylean' ancestors to be a good myth.
- Do we have evidence for this? What matters if it isn't true?
- Brandom has objected that we can't understand descriptive predicates without understanding normative ones. Is this a telling objection?
- Smith has objected that the method yields massive over-realisation. Is he right?

---

[3] See "Reduction of Mind" and "Naming the Colours" for his current view.