1 Introduction

I am concerned most directly with moral reasoning that people do. Logic, probability theory, and decision theory are in the first instance theories about consistency and implication and are not in the first instance theories about what people can or should do; so they cannot by themselves be theories of reasoning of the sort I am interested in. Indeed, it can be especially misleading to try to understand moral reasoning solely in terms of logic, probability theory, and decision theory.

I will assume that the kind of reasoning people do involves a kind of “constraint satisfaction” aimed at getting closer to a “reflective equilibrium.” Given that initial picture of reasoning, we can begin to think about what moral reasoning is.

But let me begin with some distinctions.
2 Some Distinctions

The most important distinction is the metaphysical one between the activity of reasoning, as something that that people do, and the abstract structures of proof or “arguments” that are the subject matter of formal logic. I will be mainly concerned with reasoning of the first sort, reasoning that people do. But I will also say something about the relevance of abstract arguments to the reasoning that people do.

Second, within the category of reasoning that people do, there is a distinction between moral reasoning with other people and moral reasoning by and for yourself. Moral reasoning with others may involve discussion with them, bargaining with them, and possibly arguing with them. I will be mainly concerned with reasoning by and for yourself, but I will say something about reasoning with others also.

Third, within the category of reasoning that people do, there is a distinction between theoretical reasoning and practical reasoning. Here “theoretical” is a technical term for reasoning that often changes your beliefs. (I say this is “often” so because this sort of reasoning may lead you not to change what you believes.) Theoretical moral reasoning might be concerned with possible changes to your moral beliefs, beliefs about what you or someone else ought morally to do, or what is or would be morally right, good or bad, just or unjust, fair or unfair, honest or dishonest, nice or not nice, etc.

There is also practical reasoning about what to do or not to do. In the first instance practical reasoning might add to or change your decisions, intentions, or planning about what to do or what not to do. To the extent
that practical reasoning in this sense is influenced by moral considerations, it is *practical moral reasoning*.

Fourth, we can distinguish two kinds of inference or reasoning: rapid, intuitive, mostly unconscious “system 1” inference and slower, considered, conscious “system 2” reasoning.\(^1\) The conclusion of system 1 inference may at least partly conscious, although it may have a largely unconscious structure, as in the understanding and production of language. As I will later suggest, there also may be unconscious moral reasoning with others, involving implicit negotiation or moral bargaining.

### 3 Reasoning as Change in View

Let me re-emphasize that I am mainly concerned with reasoning that consists in what I will call (potential) change in view. Such reasoning is a kind of activity that makes additions and/or subtractions to your view or (perhaps rarely) leaves your view unchanged. I will understand your “view” very broadly, so that it includes not only beliefs, but also decisions, plans, and intentions, as well as perceptions, sensations, feelings, emotions, hopes, and desires. I am mainly concerned with the activity of reasoning that usually makes additions and/or subtractions to your view (but sometimes leaves your view unchanged). I am interested in other things called reasoning only to the extent that they are relevant to this sort of change in view.

A normative theory of such reasoning is (or would be) a theory of when reasoned changes are reasonable. A descriptive theory of reasoning is (or

\(^1\)As discussed, for example in Daniel Kahneman, *Thinking, Fast and Slow* Farrar, Straus and Giroux, 2011.
would be) a theory explaining how people actually do or do not change their views through reasoning in this sense.

### 3.1 Reflective Equilibrium

One more or less familiar philosophical theory says that theoretical reasoning of this sort seeks (or ought to seek) a *reflective equilibrium* in which no further changes in your view are called for. Nelson Goodman discussed the special case of scientific reasoning.\(^2\) John Rawls, who introduced the phrase “reflective equilibrium,” was particularly concerned with such reasoning about justice.\(^3\) The general idea is applicable to all moral reasoning, and indeed to all reasoning of the sort I am calling reasoned change in view.

The reflective equilibrium approach assumes that the relevant kind of theoretical reasoning involves trying to find a balance between two or more considerations. One such consideration is *conservatism*. You seek to minimize changes in your view subject to other considerations. Another consideration is *coherence*. You seek changes that increase the coherence of your overall view by reducing negative coherence and increasing positive coherence. Negative coherence or incoherence includes inconsistency in your view and perhaps other things. Positive coherence includes simplicity and explanatory and other connections among aspects of your view. Conservatism and coherence-seeking compete with each other in that you tend to (or should tend to) violate conservatism by changing your view only to the extent that there is a sufficient gain in overall coherence. Reasoning of this

---


sort can be modeled in a constraint satisfaction network.\textsuperscript{4}

4 Practical Reasoning as Theoretical Reasoning?

It might be objected that practical reasoning can be treated as a special case of theoretical reasoning concerned with what you ought to do. On the other hand, you might also conclude that, although you ought to do A, you won’t. So a conclusion about what you ought to do is not by itself a conclusion of practical reasoning in the relevant sense.\textsuperscript{5}

Furthermore, desires legitimately play a role in practical reasoning that they cannot legitimately play in theoretical reasoning. A certain sort of wishful thinking is permissible in practical reasoning but not in theoretical reasoning. Wanting something to occur can be a reason to try to make it to occur, but not in the same way a reason to believe that it has occurred.

Another significant difference between theoretical and practical reasoning is that arbitrary choices are often allowed and even required in practical reasoning but not in theoretical reasoning. Suppose Mary needs to get to Philadelphia and is trying to decide how to drive there from Princeton. As far as she is concerned, there seem to be two equally good ways to go, one by way of route I-95, the other by way of route I-295. Having no reason to choose one of these route over the other, she arbitrarily decides on one of them and sets off on her journey. This can be a reasonable instance of

\textsuperscript{4}E.g., Paul Thagard, “Ethical Coherence,” \textit{Philosophical Psychology}, 11 (1999), 405-422

\textsuperscript{5}But what about the conclusion that you must do A? Can you conclude that you must do A but won’t? That might seem inconsistent. But, if you do not accept morality, you might conclude that morally you must do A, but you won’t, just as you might conclude that legally you must do A, but you won’t.
practical reasoning.

An hour later, Sally engages in theoretical reasoning about which route Mary has taken. Having no reason to think Mary has taken the one route or the other, it would not be reasonable for Sally arbitrarily to decide that Mary has taken the route that goes by way of route I-95? That would not be a reasonable instance of theoretical reasoning. Sally should suspend judgment about which equally good route Mary took.

Mary should choose to take one of the routes but Sally should suspend judgment about which route Mary took. Arbitrary choice is often reasonable and required in practical reasoning but unreasonable and not allowed in theoretical reasoning.

So, there are at least three respects in which it would seem that practical reasoning cannot be reduced to theoretical reasoning about what you ought to do: (1) you might conclude that you ought to do A but you won’t, (2) wishful thinking is acceptable in practical but not in theoretical reasoning, and (3) arbitrary choice is relevant in practical reasoning in a way it is not relevant in theoretical reasoning.

A possible reply to this argument against reducing practical reasoning to theoretical reasoning might go as follows.

The only sort of practical reasoning there is is reasoning that might change your beliefs about what you ought to do A, may do B, or must do C. Your desires are relevant to such conclusions only through your beliefs about what you desire. The step from such conclusions to deciding to do is not a step of reasoning
but of decision. Similarly, you may decide arbitrarily among alternatives if that you think you ought to do one of them and none is preferable to the others, but again that decision is not part of your reasoning.

Notice that complex planning typically involves several arbitrary decisions. According to the suggested reply, for each such decision, your belief that you have made that decision is an input to further theoretical reasoning about what you ought to do. What I am treating as an example of complex practical reasoning is in the suggested reply treated as a series of instances of theoretical reasoning about what you ought to decide, interspersed with decisions to do things, with the resulting beliefs that you will do those things serving as additional input to further theoretical reasoning about what you ought to decide to do.

What about revisions in your plans? How might that work according to the suggested reply? This way: You reach the theoretical conclusion that you ought to change your plans in various ways and then you change your plans more or less in the way you think you ought to.

Notice now that the reply’s attempt to make practical reasoning a special case of theoretical reasoning requires that desires influence your practical reasoning only to the extent that you have beliefs about them. Yet it seems that desires can be inputs to your reasoning without your having beliefs about them, just as your beliefs can be such inputs whether or not you have relevant beliefs about your beliefs.

In any case, I am going to continue to treat practical reasoning as dis-
tinct from, but related to, theoretical reasoning. There clearly is such a
ting as revising your plans, decisions, goals, and intentions in addition to
your beliefs. This sort of practical revision is subject to various constraints
including conservatism and coherence along with other factors, including
satisfaction of your desires and the need for arbitrary decisions.

5 Feelings

The inputs and outputs of reasoned change in view include not only simple
beliefs, intentions, and desires, but also perceptions and feelings, including
sensations and emotions.

Perceptions have content representing the environment. They are typi-
cally fleeting and have effects by directory resulting in or constituting short
term beliefs and by leaving traces in memory.

Sensations and other feelings also have content and may also involve
something like desire. The content of a pain sensation in your foot might be
to the effect that something bad is happening in your foot and might also
include something like a desire that it stop.

Hume famously suggested that the awareness of another person’s pains
may consist in your having a painful feeling located in the perception of
the other person.\(^6\) Such pain perception might contain the desire that it
cease and so function both as a belief about the other person and a desire
concerning that person.

7-8.
Antonio Damasio also offers evidence that the perception of other people’s feelings may depend on feelings of your own, so that not being able to have those feelings prevents you from responding appropriately to others and even to your own self-interest.\textsuperscript{7}

Feelings and emotions can be relevant to moral reasoning. Feeling moral disgust at a certain prospect can involve a negative moral judgment that the prospect is morally disgusting along with an aversion to that prospect. Such disgust might be an input to moral reasoning or an output.\textsuperscript{8}

Perhaps all feelings combine a belief-like aspect and a desire-like aspects. If so, feelings might be relevant to moral reasoning in the way that beliefs and desires are relevant.

6 Abstract Structured Arguments

An argument in the abstract sense typically consists in “premises,” “intermediate steps,” and a “final conclusion.” These are premises, intermediate steps, and final conclusion of the abstract argument, not necessarily premises or conclusions of someone presenting the argument. Sometimes, an argument of this sort is presented as showing how its premises imply and so support its conclusion. At other times it might be presented as an invalid argument whose premises do not imply or support its conclusion!

Of course, even when reasoning or inference involves the acceptance of an argument of this abstract sort as a good argument, the temporal order


\textsuperscript{8}There is a useful discussion of disgust in Dan Kelly, Yuck!: The Nature and Moral Significance of Disgust, Cambridge MA: MIT Press, 2011
of the reasoning may be different from the order of items in the associated abstract argument. The argument may be constructed by starting with the desired conclusion and “reasoning backward” toward the premises. Furthermore, and perhaps more importantly, the conclusion of the reasoning is not always the same as the conclusion of the abstract argument. For example, in “inference to the best explanation,” the conclusion of the inference may be one of the premises of an abstract argument.

On the one hand, I want to distinguish abstract arguments of the sort studied in logic and mathematics from activities of reasoning and inference. On the other hand, I want to try to say how inference or reasoning can involve construction or consideration of abstract arguments. In the case of inference to the best explanation the conclusion of the inference is a premise in the argument. There are, of course, also other cases in which the conclusion of an inference is a conclusion of the argument being considered.

But we need to be careful about terminology.

6.1 “Logic,” “inference,” and “reasoning”

The word “logic” is often used to refer to the theory of valid argument, taking an argument to be an abstract structure of sentences or propositions rather than an event of arguing. The premises of a valid argument logically imply the conclusion of the argument. In other words, the truth of the premises would logically guarantee the truth of the conclusion. A logic in this sense might be described by specifying axioms and rules of implication.

Logicians often refer misleadingly to such rules of implication as “rules of inference.” This is deeply misleading because these rules are not in the
first instance rules that anyone could follow. And they are not rules about inference. They are rules about when certain implications hold.

There is also another sense of “logic” to include something like a theory of method or reasoning, as in John Dewey’s *Logic: The Theory of Inquiry* and John Stuart Mill’s *A System of Logic, Ratiocinative and Inductive*. But only confusion can result from confusing these the two senses of “logic”—theory of method and theory of implication.

Defenses of nonstandard logics are typically derived from such a confusion about rules of logic: relevance logic, logics of entailment, paratactic logics, and variants seem to be based on something like the following line of thought.

In standard classical logic, any arbitrary proposition follows from contradictory premises. But it would not be rational to infer from a newly discovered contradiction in one’s beliefs to any arbitrary proposition. So, we need a logic in which it is not the case that an arbitrary proposition follows from contradictory premises.

Similarly, the idea that we need a logic of practical reasoning, such as a logic of imperatives seems to rest on the following confused thought. “There is practical reasoning, so we need a logic of practical reasoning, perhaps a logic of imperatives.” And the idea that there must be a logic of induction seems to rest on a similar confused thought. “There is inductive reasoning,

---

9Holt, Rinehart and Winston: NY, 1938
10London: J. W. Parker, 1846
so we need an inductive logic.”

6.2 Logic is not a theory of justification

Deductive logic is a theory of implication and consistency. It is not by itself directly either a normative or psychological theory. It is not directly a theory about what to believe or what to do. It is not a theory about what justifies what.

A rule like modus ponens is a rule about implication. It says that a proposition $P$ and a corresponding proposition if $P$, $Q$ together imply $Q$. It does not say that, if you believe $P$ and also believe if $P$, $Q$, you will, may, or must believe $Q$. Deductive logic says nothing about belief or inference either descriptively or normatively.

6.3 How Important Is the Ability to Produce Reasons for Your View if Challenged?

Hugo Mercier observes that people are motivated to produce reasons for their views if challenged. “This is important as, if the rationalizations are voiced out, others can then attack them. If they are found lacking, if we fail to defend our original judgment appropriately, then we must change our mind on pain of being thought to be inconsistent and unreasonable.”

Similarly Paxton and Greene say, “People reject judgments based on their own intuitions when those judgments appear to be unprincipled . . .”

---

Now, it is true that people are often willing to produce reasons for their views, but it is well known that these are often invented on the spot and, if the reasons are shown to be inadequate, people typically come up with other reasons.

Furthermore, people do not typically reject their judgments just because they cannot justify them, as is illustrated in Jonathan Haidt’s “dumbounding” experiments. For example, his subjects tended to be quite certain that incest is wrong, even after all their reasons for thinking it wrong were shown to be inadequate.14

7 Believing the Consequences of Your Views

A related point is that it is not generally true that you ought to believe the logical consequences of your beliefs. For one thing, there are infinitely many distinct logical consequences, and you cannot believe them all. For another thing, your beliefs are almost certainly inconsistent, which means everything follows from them, and you absolutely should not try to believe everything.

Discovering an inconsistency in your beliefs does not require you to drop everything in order to resolve it. You have other concerns. Perhaps you are trying to find your wallet. Or you may want to wait until after dinner or until after the latest episode of a TV series you have been watching.

Even if nothing else is urgent, you may not care enough to try to resolve an inconsistency. Given a short apparent proof that $1 = 2$, most people

have little or no interest in trying to figure out what is wrong with the proof. Similarly, they may be convinced that their views about sets are inconsistent, but simply not care.

Even if you care, you may not know how to resolve inconsistency. If you come to appreciate the liar paradox and would like very much to resolve it, you may find that your beliefs about truth are inconsistent and you have no idea how to avoid inconsistency. Even experts have trouble with this.

In the 19th Century, Kelvin’s calculation of the age of the earth using principles of thermodynamics gave a result that was too small to allow for what was calculated to be needed for evolution. It seemed then that the theory of evolution was inconsistent with thermodynamics. Kelvin took this to disprove evolution by natural selection. But many theorists continued to accept both evolution by natural selection and thermodynamics, despite their apparent inconsistency. This was not resolved until the discovery of atomic radioactivity, a previously unknown source of energy that Kelvin had not allowed for.  

It is not easy to say what connection if any there is between deductive logic and what it is rational or reasonable for you to believe.

If your beliefs immediately and obviously imply $P$ and you have a good reason to be interested in whether $P$, that can give you a disposition to believe $P$ in the absence of any reason not to. But notice that if your beliefs are inconsistent, your beliefs may immediately and obviously imply not-$P$ too.

How to change your beliefs after recognizing an implication of them or

---

an inconsistency in them can sometimes be a complex and difficult issue.
At other times, you just immediately react by accepting an implication or
abandoning one of the inconsistent beliefs.

8 Sketch of a larger role for abstract arguments in reasoning

Some theorists might grant that, at least in the first instance, logical rules
are rules of implication rather than rules of inference, but might go on to
say something like this:

Reasoning always consists in the construction of an argument
for a conclusion. A relevant argument might be more or less ex-
licitly deductive, but it might also be inductive. It might be a
default argument, whose premises do not guarantee its conclu-
sion but only make it likely. Such an argument might contain
another argument as a part, as in an inference to the best ex-
planation that contains an explanatory argument. It might in-
volve “negation as failure”—you accept a certain generalization
because you have not been able to find a counter instance. It
might be a practical argument, whose conclusion is a decision to
do something.

So far this makes it appear that reasoning always adds things to what
you accept. What about the ways in which reasoning can lead you to revise
your beliefs and plans in part by giving some of them up?
When there is an argument for giving up certain of your beliefs, accepting that argument involves giving those beliefs up. Similarly, if you come up with an argument for changing your plans in certain ways, accepting that argument involves making that change.

In fact, whatever are the correct principles of change in view, in any particular case there will be a corresponding *argument* for making the relevant changes to your view: “the correct principles of change in view imply that, given your current view, such and such changes should be made, so you ought to make those changes.”

But now the theory that reasoning always involves the construction of an argument is plausible only because that theory is parasitic on a theory of reasoned change in view.

9 Reasoning as Constraint Satisfaction

I mentioned earlier Nelson Goodman’s suggestion that we take beliefs to be justified to the extent that they cohere with other beliefs. I noted Rawls’ use of the term “reflective equilibrium” as the goal of such reasoning. Paul Thagard and colleagues have programmed computers to run complex constraint satisfaction networks that take emotions and beliefs into account.¹⁶

Constraint satisfaction networks have been used to model certain aspects

of the visual experience of a Necker cube. One tends to see the image not just as a collection of lines but as a cube with either one or another face in front. Thagard’s models tend to stabilize much as the appearance of a Necker cube does. In particular, Thagard’s models are sensitive to the order in which information comes in.

One of Thagard’s models has been used to simulate a juror trying to reach a verdict. The fit seems pretty good. In particular, as a juror becomes confident of a particular verdict, the juror’s confidence in principles that support that verdict go up and the juror’s confidence in principles that speak against that verdict go down.¹⁷

This means that verdicts can be unstable in the sense that small changes in the presentation of evidence can have large effects on the verdict. In particular, the order in which evidence is presented is quite important in this way.

The method of reflective equilibrium may be unstable in the same way. Similar considerations have been used to argue that conclusions about justice should be reached via taking a “wide reflective equilibrium” rather than a “narrow reflective equilibrium.”¹⁸

All this suggests that much of the work in reaching reflective equilibrium takes place unconsciously in the network. In wide reflective equilibrium one would have to try to reason from different starting points. This might be encouraged through sincere discussion and attempts to understand others.

who have different starting points.

10 Moral Bargaining

Rawls’ theory of justice is indebted to social contract theories as exemplified in the writings of Hobbes, John Locke, David Hume, Adam Smith, and others. Rawls’ is concerned with a hypothetical contract meeting certain conditions. Following Hume, we might consider the idea that there can also be an actual ongoing but often implicit social contract or set of conventions among members of a given group. The relevant conventions may be implicit and difficult explicitly to specify. Instead they would be reflected in how members of the group act and expect others in the group to act.

The conventions existing in a particular society may be the result of implicit social bargaining and adjustment, similar to conventions about what side of the sidewalk to walk on. Hume offers the example of rowers who end up rowing at the same rate as each adjusts his or her stroke to that of the others. Each rower typically would not be able to specify the rate of rowing except to say that it is the same as the others.\textsuperscript{19}

There might be a moral convention of noninjury among members of the group as well as a convention of mutual aid. If some participants have more power and wealth than others, it might be expected that the convention of noninjury, which benefits everyone, has more force than the convention of mutual aid, which mainly benefits the poorer and weaker members of the group.

\textsuperscript{19}David Hume, \textit{Treatise on Human Nature}, Book 3, Part 2, Section 2.
The moral conventions that actually exist in most social groups appear to fit better with such actual (as opposed to hypothetical) social contract theories than with utilitarianism.

The sort of adjustments involved in rowing and in getting along with others are possible examples of unconscious reasoning with others I mentioned earlier, involving a kind of implicit negotiation or moral bargaining.

11 Final Summary

I began with four distinctions: (1) between reasoning as something people do and abstract structures of proof, (2) between reasoning by oneself and reasoning with others, (3) between theoretical and practical reasoning, and (4) between rapid unconscious system 1 reasoning and conscious system 2 reasoning.

I then said I was mainly interested in the sort of reasoning involved in change in view. I distinguished theoretical reasoning of this sort from practical reasoning and argued against trying to reduce one sort to the other. In either case, I suggested that reasoning seeks a reflective equilibrium.

I briefly mentioned how feelings can be relevant to reasoning by serving as a mixture of belief and desire.

I noted ambiguities in words like “argument,” “reasoning,” and “inference,” between processes and abstract structures. I noted that the premises and conclusions of an abstract argument are not necessarily premises or conclusions of someone presenting or reasoning about that argument. I stressed the importance of distinguishing issues about abstract arguments as stud-
ied in logic and mathematics from issues about processes of reasoning. I was doubtful about the importance of being able to produce reasons if challenged.

I observed that you cannot believe all the consequences of your view and that in any event your views are almost certainly inconsistent.

I discussed how reasoning might be best conceived as attempting to satisfy a variety of constraints like conservatism and coherence. Thagard’s models of this sort of reasoning exhibit a certain instability. Attempts to reach wide reflective equilibrium might do better.

Along the way, I suggested that treating all reasoning as involving argument construction will work only if such a treatment is parasitic on an adequate theory of reasoned change in view. More generally, the policy of developing a variety of logics to account for reasoned change in view is a degenerating research strategy.

Finally, I suggested that morality might be the result of implicit bargaining among people with different powers and resources.