Davidsonian Rationality and Ethical Disagreement between Cultures

In *Davidson's Philosophy and Chinese Philosophy: Constructive Engagement*, edited by Bo Mou, Brill, 2006, pp. 165-188.

Introduction:

Chinese ethical thought has long been a paradigm for the West of a way of thinking that is sophisticated, highly developed, advanced on any reasonable scale of advancement, but very different in many judgments from the ethical thinking of the West. Stock examples include differing degrees of respect for authority, different views about the responsibility of groups for the actions of individuals, and the like. This essay investigates whether such divergence shows that ethics is less "objective" than natural science. This essay concerns the abstract question of the objective truth-values of ethical sentences, rather than any particular difference that might be claimed to hold between these two cultures' ethical thinking.

I Kant and Founding Ethics on Rationality

Many ethical theories have attempted to show that it is rational to do the good. The idea is that the relatively uncontroversial normativity of correct reasoning can found the normativity of morality. The most ambitious of these theories was that of Kant. Kant attempted to show that it was irrational to do anything other than the right thing. The argument for the categorical imperative is an argument that the very concept of acting on purpose requires that a perfectly rational agent do only what has a coherent general principle. The least a perfectly rational agent could do, given that he was acting for reasons, would be to act so that he was consistent, because acting for reasons requires acting according to considerations that apply generally. That is, if the considerations are universally applicable, it ought to be possible that everyone follow them. Thus acting for reasons consistently requires acting in a way that anyone could act.

Kant hoped that this criterion of consistency would suffice to give both necessary and sufficient conditions for moral actions. Kant’s program was quite ambitious. He not only tried to supply a rational ground for doing good when a person knows what the good is, he also required that there be a rational basis by which one could always know what the good is. Rationality, that is, would not only provide a reason for doing the good; it would also provide a criterion for the good which would enable a person to be free from sin, i.e. which would enable a careful person always to do the right thing.

Kant’s conception of “doing the right thing” is “doing one’s duty” or “fulfilling one’s obligations”. An ethics based on obligations and duties is essentially deductive, since principles of obligation yield arguments that can be known to be sound on the basis of limited information, i.e. yield deductive arguments. Thus, a Kantian theory requires that ethics have the structure of a system of principles, so that ethical reasoning is primarily deductive. Given adequate principles, a person is equipped to know the right thing to do in every circumstance. Thus the person must be able to deduce the correct thing to do on limited information.

II Davidsonian Kantianism in outline

A combination of several ideas in Donald Davidson’s philosophy suggests a way to reconstitute something like Kant’s project. For Davidson, roughly, the normativity implicit in reasonable interpretation of another as an agent supplies the basis for understanding the normativity of ethical concepts. Interpretation, though, uses a much richer notion of rationality, one that includes “inductive” logic. The conception of ethics that results is not deductive. No algorithm, for instance, will allow even a perfect rational agent to always do the right thing. Davidson’s ideas allow us to connect the normativity of rational interpretation to the normativity in the truth-conditions of ethical language. On this account, the normative is not different in ontological or epistemic status from the "factual".

Let me sketch some of these Davidsonian ideas:

a) Interpretation as Maximization of Agreement:

Davidson, following the suggestions of Quine, uses the relatively testable and empirical notion of what we do in understanding another in order to unpack the concept of "rational agent". The idea is to determine the content of "rational agent" or "entity acting for reasons" by seeing what constraints we apply in interpreting another.[1] Davidson's account of interpretation yields an account of rationality that is rich enough to give hope that one could show that doing good is reasonable. Interpretation maximizes agreement, so that interpretation is constrained by the necessity for treating the other as mostly believing the true and wanting the good. "Maximize agreement" is a formula for a set of probabilistic constraints whose content is, roughly, "Interpret on the supposition that the entity being interpreted is an agent." "Maximize agreement" says, "Interpret the agent by using the features of a paradigm agent, namely oneself." To interpret according to a probabilistic fit with our own case of rational agency is essentially to apply a term on the basis of a hypothesis that the other is like us, an agent. Thus, since this is the general procedure for applying predicates, the reasonableness of such constraints is a priori.

Familiar examples of the operation of maximization of agreement discuss maximization of agreement in beliefs. But action interpretation must also maximize agreement in desires, that is, conceptions of the good, as well. When someone drops a rock on his foot, the interpretation that the person wanted pain and believed this was an effective way of bringing it about is reasonably rejected, barring very special circumstances,[2] and interpreted as unintentional. Action interpretation always maximizes agreement in both desire and belief. Given only that the other entity is an agent, and given only that x is a belief or desire, if we have that belief or desire x, so, probably, does the other entity.[3]

A broadened conception of the rational that constrains the content of desires is required for action interpretation. Without such constraints on content, interpretation could not get started, since any behavior is consistent with any beliefs, if that any desire whatsoever is as likely to be present as any other.[4] So, such constraints as that a person ought not to want pain are part of the concept "is a rational agent".

Speech interpretation is a special case of action interpretation—what a person is doing in speaking. In detail, what a person is doing is, for example, asserting that the cat is on the mat. Speech interpretation, though, requires hypotheses about the speaker's intention in order to get started. For instance, in interpreting an utterance as a sincere assertion, one applies "maximize agreement in beliefs" in the light of the hypothesis that the person desires to say the truth. Thus we seek a truth-definition for the language the person is speaking that will make the assertion true. Given that the utterance is a sincere assertion, i.e. that the intention is to tell me what is the case, and given that the speaker is by and large a believer of truths, the utterance ought to be true. The "ought," which is close to a "probably," falls out of the very structure of the constraints on interpretation—that they are maximization constraints and not absolute constraints. Briefly, principles of interpretation are "ought" principles.

b) Ought sentences

Davidson's discussion of weakness of the will[5] presents a brief account of conditional "ought"-sentences (hypothetical imperatives) that suggests a connection of ethics with interpretation-theory. Roughly, conditional "ought" sentences are formally akin to conditional probability sentences, in that they do not "detach" when the antecedent is true. In logical form, conditional “ought”-sentences are relative to something like an consideration-base, akin to an evidence-base to which a conditional probability claim is relative.

That is, the apparent “antecedent”[6] of a conditional probability claim, an “If A, then probably B” claim, is really the evidence relative to which the “consequent” is probable.[7] [8] By analogy, conditional “ought” sentences give the bearing of a consideration on whether a person ought to do something. “Ought,” like “probably,” is fundamentally inductive.

In contrast, “obligation” sentences, the focus of Kant’s ethical theory, have the logic of deductive argument, as outlined in various deontic logics. It is arguable that “ought”, rather than “obligation”, is the primary ethical concept, since one can always ask whether one ought to fulfill an obligation. If “ought” is the primary ethical concept, and Davidson is right about the form of “ought”-sentences, then we get a very different picture from Kant’s of the connection between rationality and ethics. Ethical arguments and ethical thinking are most properly understood on the model of induction rather than deduction. In the typical inductive argument, a set of true premises {p1…..pn} which strongly leads to the conclusion B is compatible with the existence of a true premise pn+1 which, together with {p1…pn}, strongly leads to the conclusion not-B. In an ethical consideration, a set of truths{p1…..pn} may strongly lead to the conclusion that you ought to do B, while there can be another truth pn+1 which, together with {p1…pn}, strongly leads to the conclusion you ought not to do B.[9]

If ethical argument is like induction, then ethical reasoning does not proceed according to Kantian universal moral principles. Briefly, an ethical actor will not in general know that he is doing the right thing, since he is always operating on limited information, using guidelines that allow for the possibility that further information can undermine the conclusions so far reached. The “principles” that can be applied are not universally quantified commands, but rather “rules of thumb". No algorithm will determine, given a description of a situation, what in fact a person should do in that situation. Furthermore, the rationality that might be the foundation of ethics would require much more than formal consistency. Ethical truths will be much more like scientific truths than logical truths. Just as there is no algorithm for the construction of a new theory in the light of novel experience, so there will be no algorithm for the correct ethical judgment, given that every situation is novel in some respects.

Davidson's account on which "ought"-sentences have a form similar to that of conditional probability sentences suggests a theory of "ought"-sentences. If "ought"-sentences are essentially like conditional probability sentences, then their categorical form has an implicit relativization. Just as you may say "It will probably rain," so one may say, "We ought to help those people." In both cases, there would be an implicit relativization, represented more explicitly by an "All things considered" clause.[10]

Such relativizations suggest a unified account of the distinctions between the various "senses" of "ought" that thinkers have distinguished. Consider the general principle that one ought to believe the logical consequences of what one believes. This principle seems to have obvious counter-examples. If Fred believes that Susan is honest and that Susan has been embezzling his funds, he should not conclude that honest women sometimes embezzle, but rather ought to give up one of his premises. Likewise, if the consequences of a set of beliefs include some beliefs that are illegal, and if an inquisition can reliably find out that you have such an illegal belief, then you ought to abandon some of your beliefs that lead to the illegal consequences.

The idea is that the various "senses" of "ought" are to be understood as further relativizations to various backgrounds. For "ought"'s applied to human intentional actions, various implicit "consideration" relativizations yield different "ought"'s. The "prudential" "ought" might be relativized to an implicit “set” of a person’s self-concerned interests, the “logical” “ought” might be relativized to considerations of consistency, and so on.

A Davidsonian account should attempt to treat all “ought” sentences as having the same basic structure and semantics. The methodological principle operative here is that, other things being equal, a theory should minimize homonyms. It could strike someone as implausible that the “ought” in “You ought not to inflict pain” and in “If you turn the key, the car ought to start” are the same word. Indeed, the two occurrences may be different words. However, a plausible theory that makes the two “oughts” the same is preferable to a theory that multiplies homonyms.

III Relativizations

a) the two kinds of relativizations

We need to state carefully what exactly these relativizations consist of. A conditional “ought” utterance, just like a conditional probability utterance, has two general kinds of relativities. These are further complicated by the fact that they are sometimes implicit rather than overt:

1) The first kind of relativity is often overt, and is specified in the “antecedent” of the utterance. So, “If you want a nice dinner, you should go to Cavey’s” recommends an action relative to a desire. On the normal understanding of such an utterance, the utterance is a “prudential” ought. The “prudence” indicated is relative to having the desire specified in the antecedent. Given that desire, and given information that you have, and so that the other should have, the action that the other should (prudentially) take is to go to Cavey’s. When the utterance is a prudential “ought,” the antecedent gives a special circumstance which, given “prudential” considerations, make it reasonable for a person to choose Cavey’s.

These “conditional” relativizations describe what, relative to the “sense” of “ought” (the “considerations” described below), a person should do given all the surrounding circumstances.

That is, in the light of all the information available,[11] if you want a good dinner, you should go to Cavey's. The bearing of the information on what you should do is by and large inductive. Features of Cavey's, information about your tastes, information about alternatives, and so forth reasonably lead to that conclusion, albeit not by an algorithm. In effect, such reasoning is a "practical syllogism" with the "syllogism" expanded to include more adequate forms of reasoning.

2) The other relativity is what distinguishes the various "senses" of "ought". This is relativity to considerations intended to be taken to bear on the choice.[12] To illustrate, consider a player in a chess game, in a position where she can force mate in four with a bishop sacrifice at f7, while every other move lose. In one "sense" of "ought," it is clear that she should go bishop f7 check. But her opponent is her boss, who will take it seriously amiss if this young employee defeats him in chess. So, prudentially bearing in mind her own personal welfare, she should make some other move than bishop to f7 check. But her loss to the boss may demoralize her colleagues to the extent that their lives are significantly worse if she loses, even though her own life is worse if she wins. Or, it might be dishonorable to lose a game on purpose. The noble thing for her to do, and perhaps what she ought to do, is win.[13]

These different understandings are different restrictions on the considerations relevant to what to do. In the first case, the restriction might be called "chessic". Only the goal of chess, mating the opponent's king, is taken to be relevant. What the person ought to do, bishop to f7 check, is taken relative to just those considerations. Considerations that concern only her personal welfare would yield a prudential "ought". If the recommendation that she move bishop to f7 were interpreted prudentially, the claim would be that, taking into account only considerations relevant to the "self-interested"[14] desires of the agent, she should go bishop to f7.

The sentence itself does not determine this relativization. Rather the intent of the utterance does. So, even though the unqualified recommendation in a book of chess problems that bishop to f7 is the best move, it may not be, given another intended set of considerations.[15]

Such restricted background considerations can be regarded as a set of interpreted sentences,[16] but the conditions of set membership are too vague to make that a precise proposal. In particular, the notion of "personal welfare" and "self-interest" will turn out to be irredeemably vague. The background considerations will have to remain an unanalyzed primitive, akin to the backgrounds that are presupposed in remarks that it will probably rain, or in counterfactuals.

IV Questions of logical form: How do these two kinds of relativizations work?

As Davidson showed long ago, the relativizations cannot be thought of as having the form of conditionals that detach.[17] Thus the form of a conditional "ought"-sentence cannot be any kind of connective joining two independent clauses, one of which has an "ought". This strongly implies that the conditional "ought" is the basic notion, and that the absolute, categorical "ought" is something like "ought, all things considered. This accords with recent results in probability theory that argue that conditional probability, rather than absolute probability, is the fundamental probability notion.[18] So, an "ought"-sentence must be a two-place modality, a construction that depends on more than the truth-values of the two clauses. Modality brings in a complicated set of difficulties about logical form and ontology on its own.

A difficulty, then, is that a serious Davidsonian account of the form of "ought"-sentences will have to include a general account of modalities, propositions (or their surrogates), and conditionals generally. This is surely a daunting task, and one that is not by any means complete. Let me sketch some of the strategies with which this Davidsonian is working:

An account of the form of “ought” sentences must accommodate several constraints, none of which will be addressed in any detail in this essay:

First, the account should treat “if..then” in a way that generalizes. “If” must be doing the same thing in “if..then..ought” sentences as in other occurrences.

Second, some account must be given of the objects that take the place of A, B and C. If “ought” is a predicate, the considerations C, the “antecedent” A, and the “consequent” B must be construed as some kind of entities. As a paleo-Davidsonian, I would assimilate such objects to the general type of demonstrable linguistic objects, along the lines of “On Saying That”.

Third, some general account of modalities will be required. For a Davidsonian, it is arguable that modality must be primitive.[19] The different “senses” of the usual modal “operators”[20] will be different relativizations to considerations, as sketched above for “ought”. “Considerations,” on this account, are construed as sets of “things said” that, for the case of “can’t,” function as limitations on the principles from which the negation of the sentence is a consequence.[21]

The modalities typically treated, “necessary,” “possible,” and the like, seem to be based on logical consequence, and so would count modal operators generally as having deductive underpinnings, relative to different considerations. The paragraph above is thus really a re-statement of ideas that Carnap and others had in the last century.

Given that there is no hope of reducing induction to deduction, analogous modalities that depend on inductive connections would use a primitive “consequence” relation about which there is only the beginning of a theory. I would argue that “If..then.. probably” and “if…then..ought” express such modalities.

So, my working proto-theory is that the form of conditional “ought”-sentences is Ought (C, A, B) rather than (A->Ought_C B), where “C” stands for “considerations,” and where “Ought” is a modality whose rough content is that “Ought(C, A, B)” is true if and only if, given A, B is reasonable on considerations C. That is, rather than a conditional “ought” sentence being some kind of conditional with a consequent containing a “sense” of “ought,” such a sentence is rather a three-place modal predication among an “antecedent,” an set of considerations, and a “consequent”. In order to be objects that fill places in a predication, A and B and C must be some surrogate for propositions or sets of propositions, rather than names of truth-values. Thus “ought” is to be treated as a predicate, with the “relativizations” specified by A and C and the action by B.

V Connections between “ought” and “conditional probability”

“Ought”’s often amount to the same thing as “probably”. “If you turn the key, your car ought to start” is usually true if and only if “If you turn the key, your car will probably start.” The relativization in a probability sentence is not merely to the contents of the “if”-clause, but to what might be called “background conditions” as well. The assumed overall conditions that the laws of nature obtain, the material in the tank is gasoline, and so forth would also be “implicit” in the “antecedent” or relativization of the “probability” sentence.

While “if then probably” sentences and “if then ought sentences” have a number of connections, a sharp definition of one in terms of the other now seems to me not to be likely or necessary. My current thinking is that “ought” and “probably,” while they have inter-connections, are each semantically independent modal predicates. A rough account of “ought”’s connection with “probably” would rely on the probabilistic, maximization character of agreement. Given any feature that correct interpretation maximizes, if an individual being interpreted is indeed an agent, then that individual, relative to being an agent, probably has the feature. For any truth, then, relative just to an entity being an agent and the truth being a truth, the agent probably believes it. “Ought” is then roughly a chain of such probabilistic connections. The idea is that a person who believes as he ought, given his information, makes *every* inference that he would probably make.

The picture of ethical reasoning is that any complex logical or practical inference can be broken down into a series of smaller steps. Each of the beginning premises is a desire or belief the person probably has. Each of the smaller steps is one that a person would probably make, (and so ought to make), just in virtue of being a rational agent. So, ethical reasoning would arrive at "ought"-sentences by constructing what a person would do if that person always did what, on principles of reasonable interpretation, the person would probably do.

For most real-life cases of ethical decisions to be made, the chain of calculations and inductive judgments in the ideal determination of what a person ought to do will be enormously complex. Although some actual ethical reasoning is something like the ideal reasoning, most of the time various short-cuts are appealed to. The short-cuts, or ethical principles, are rules of thumb that by and large give the right results. They are analogous to the irrationalities that Twersky[22] and others have researched. Given the information-costs of thoroughly calculating every decision, it pays to have devices that give the right result much of the time. Such short-cuts, however, on occasion give the wrong result. So, the ethical principles that Kant took to constitute the core of ethical reasoning, and which allow deductive determination of what a person ought to do on a given occasion, are not, from a Davidsonian perspective, central at all, even though they are important and largely true.[23]

VI The Moral "Ought"

a) Outline

According to this theory-sketch, the "moral" "ought" would be the minimal restriction of considerations, the relativization of relevant considerations just to "is a person" or "is a rational agent". The "moral" "ought" is "ought, all considerations considered" (considered in the light of what it is to be an agent). The rationality that is built into interpretation is not just formal, but substantive: What a person desires at bottom is interpreted by "maximize agreement". The thought is then that, given rational desires (in the content as well as structural sense) and given rational practical inferences, what a person ought to do will turn out to be the output. In effect, this yields the familiar idea that, given that people basically want the same things and are basically reasonable, there must be a basis of moral agreement. Disputes, one might hope, can be settled by reasoned discussion rather than by brute force.

For this idea to yield something like morality as we understand it, the rational desires will have to be strongly shaped by an argument that other peoples' desires rationally count, just as a person's own desires do, i.e. that it is irrational to show exclusive concern for one's self-interest. Kant tried to show this by arguing from the very notion of purposive action, and its connection with universalizability. More recently, Thomas Nagel's book, *The Possibility of Altruism*[24] tried to establish this, likewise by reflecting on what it is to do something for a reason. Philosophers have various opinions on whether these proofs work or whether demonstrations of some other sort work. It seems to me that unless a connection between rationality and counting others' desires as rationally motivating holds, the "moral ought" is a chimera. I have no arguments better than Kant's or Nagel's, although my discussion of preference below will try to make the Kant-Nagel thesis seem natural.

Even given the soundness of the Kant-Nagel argument that others' desires count, can an adequate moral theory resolve ethical issues and arrive at ethical truth? I will argue that the answer is, "Sometimes, to exactly the extent that physical discussions can resolve disagreements about `fact'." This weaker result leaves us with a substantive morality, but without the idea that every ethical question has a truth-value we can determine by any kind of calculation or discussion. Sincere people, perfectly rational from their own points of view, can have irresolvable moral disagreements. But this does not mean that there are genuine moral questions that lack truth-value. The presumption that irresolvability means relativity for ethics but not for physics seems to me to rest on a mistaken view of preferences.

The lack of resolvability will be most apparent between cultures with different histories and different traditions, as we will discuss below. However, we should expect the same irresolvability, perhaps on a smaller scale, between any two individuals. Just as for a Davidsonian the fundamental language is the idiolect, so in cultural terms the elementary culture is the individual. “Same culture,” like “same language,” has no strict sense. Nevertheless, of course, I can loosely say that I speak English and am a Westerner.

The difference from Kant’s hoped-for result derives from the nature of the constraints on interpretation that give a Davidsonian conception of rationality. One important aspect of these constraints, and the rationality that is constructed on their basis, indicate that irresolvable differences will not disappear. The constraints are maximization constraints. The other shares one’s beliefs and values only probabilistically. Even when the values are shared, the ranking of any pair of goods is only probabilistically the same. Furthermore, the inferences we are willing to make from what we believe are only probabilistically identical. So, in effect, just as each of us has his own idiolect, each of us has his own rationality.

This feature is part of the cost of a notion of rationality that goes beyond mere structure, but rejects the analytic-synthetic distinction. Allowing that there can be rational constraints on content,[25] so that an interpretation that ascribes a desire for pain is probably rejected, would still give us a sharp notion of the rational if there were a sharp line between desires that are essential to being rational and desires that are not. However, for a Davidsonian, there is no sharp line between a part of the theory that is essential and a part that is contingent. Of course, there are important differences between parts of theories, and important differences between desires. Some desires can only be ascribed given overwhelming evidence; others are ascribed failing strong evidence that the person lacks the desire. But there seem to be neither desires every rational agent has to have nor desires that no rational agent could have.[26]

b) A Problem with Objectivity?

Both of the above considerations, though, apply to physical theories as well as to opinions about what one ought to do. Why do questions of ethics appear more subject to unresolvable disagreement than questions about truth-values of non-ethical claims? The usual answer is that, whereas agreement in physical theories involves only getting beliefs in agreement, ethics involves beliefs and desires. Desires, it is claimed, unlike beliefs, do not "correspond" to anything objective. People differ in preferences about which there is nothing like a "true" or "false" to supply objectivity. Even though people by and large agree on goods, their rankings of goods often differ. Any such differences in the complex calculations that arrive at what one ought to do, make a difference. That is, since an ideal calculation of what we ought to do in any complex type of situation involves very many small steps, small differences in the desires can lead to differences in judgment.

Unlike differences in belief, differences in desires, it is held, reflect nothing but internal differences in persons. There seems to be nothing like triangulation for desires. There appears to be no common world of value to which both an interpreter and the interpretee are related to when an individual expresses a preference. So the interpretation of actions lacks one of the groundings in a common world that interpretation of belief has. There appears to be no "common world of correct desirability" by which people learn what is better than what. I argue below that this apparent lack of a common ground that functions as a common world is illusory.

c) A Quinean Hedonistic Solution to Objectivity

One might hope that physiological states such as pleasure and pain could be the stimulations that ground all desires by supplying something analogous to the common world in which belief-triangulation takes place. A Quinean version of hedonism could then say that the relation between the good and pleasure and pain is like that between external stimulations and a physical theory. Just as there is no sentence-by-sentence reduction of physical sentences to patterns of external stimulation, even though the theory does nothing but organize stimulations, so there is no good-by-good reduction of valuations to pain and pleasure. An account of valuations would thus be grounded in internal stimulations, but holistically. A "web of belief and desire" would be a conception of the "sensory meanings" of sentences that distributed degrees of belief and valuation among the sentences of the language. This would be a kind of pragmatism, since the organism would desire whatever, to the best of its knowledge, maximized its welfare. Differences in preference would be explained by differences in physiology, leading to a kind of innocuous relativism.[27]

If the arguments that it is rational to value other agents' welfare are sound, the theory would be a grounding for morality as well as for prudential thinking. Such a theory would say that, just as it is unreasonable to take one's own perceptions as the only reliable ones, so it is unreasonable not to take the preferences of others as counting in the determination of the objectively valuable. Quinean hedonism could thus be a universal utilitarianism of a sort.

The difficulty with such a theory is analogous to the difficulty with the Quinean Web of Belief. The theory supposes that sensations are given unconceptualized desires, the periphery of desire-stuff that is organized via beliefs into particular preferences for, e.g., volleyball over badminton. Absent such a given, the special tie to pleasure and pain as physiological phenomena would drop out. But something akin to the above theory could be a Davidsonian account of ethical truth.

VII Davidsonian Ethical Objectivism

A Davidsonian can treat ethics as objective, and hold indeed that every ethical sentence has a truth-value, even without a basis in pre-conceptualized desire-stuff. Consider the apparent difference between beliefs and preferences that marked a difference between ethics and physics above. We need not regard differences in preferences as brute, just as we need not treat idio-rationalities as making "rationality" relative to persons.

Just as we do not generally identify beliefs with truth, so there is no reason to identify preference with our good or with the good. Given our past experience, we know that some of our beliefs are likely to be mistaken. In the same way we find that some of our desires were ill-considered. Thus, just as we do not identify the true with our beliefs, so we do neither identify the good with what we want nor identify the rational with our total theory. We can realize that we are less than perfectly reasonable, even though we hold, of each particular view, calculation, and valuation, that it is reasonable.

The decision-theoretic tradition has taken preferences as just brute facts about which nothing is rationally required but coherence, whereas beliefs are true or false. But in interpretation, we clearly take some preferences to be irrational, such as the simple preference for pain over lack of pain. Some preferences can only be assigned to an agent on the basis of very strong evidence. "Pain is worse than no pain, other things being equal," is true. So, prima facie, some "better than" sentences are objectively true. If some “better than” sentences are true, the reasonable supposition is that “better than” sentences usually have truth-values.

The conception of preferences that suggests itself is that a preference that A rather than B is a belief that A is better than B.[28] The causes that give rise to preferences would then be construed as akin to the causes of the involuntary utterances[29] that are sensory judgments. On this conception, pain and pleasure would, in motivating action, be involuntary utterances “This is bad” and “This is good.” Such involuntary utterances, of the speaker to himself, as it were, are incorporated into the web of belief and desire. Sometimes they are over-ridden by other considerations, just as sudden apparent flashes of light are over-ridden by the consideration that you may have an eye problem. That is, we voluntarily get our separated shoulder replaced in its socket by a very painful procedure. Such over-riding is not the discovery that the pain is good, but rather that, while the pain is bad, the conjunction of the pain and the cure is good. Given that the person is reasonable, and knows that the same event will be both a pain and a cure, the person takes steps to bring about the event.[30] But this does not mean that the person’s evaluation of the pain itself was mistaken. A kind of practical failure of reason can explain how that a person cannot bring himself to bring about the event, even though it is better and, “intellectually,” known to be better. In that event, the immediate evil, the involuntary utterance that is the pain, is over-rated relative to the future good. I argue below that we can regard the wide-spread disposition to choose for actions favoring one’s own interests as another example of the same kind of failure of reason.

VIII “Better for me,” “self-interest,” and objectivity

If agreement on the Good is indeed what we maximize in interpretation, we should not take into account only our own involuntary utterances about the Good and the Bad. If triangulation is to apply to the Good, the objectively good must be prior to both what I think is good and what is good for me. Only if that is the case could we claim that everyone's desires reasonably ought to count as reasonable for me to take into account. I will make this plausible in two stages: First, I will argue that "self-interest" is by no means as clear a notion as it has often been taken to be. Second, I will explain how it is that most valuations can be correct while almost everyone fails to take the moral point of view. I will then conclude the section by examining what sort of ethics a Davidsonian will have. I will argue that we should not expect that a Davidsonian ethics will be a utilitarianism. Whether and to what extent ethical sentences have non-relative truth-values will depend on what is the case about partiality.

a) What is "Self-interest"?

In outline, the argument of this sub-section will be as follows: The particular interests of any normal agent go far beyond anything identifiable as particular states of the particular organism that is the agent, or "self-interest" as usually conceived. Usually the identification of one's interests with the interests of others is limited to a subset of the others, and so would be termed a "partiality," a discrimination. Ascribing such partialities is part of interpreting action and speech, and so part of the rationality that is part of interpretation. Interest and lack of interest in the welfare of others is thus subject to rational evaluation. "Reasonable" attachments are maximized in interpretation. But, if Kant and Nagel are right that it is rational to extend one's concerns to the interests of every entity that has interests, then it is reasonable to extend the scope of one's concerns with the interests of others to everyone.

If "X's self-interest" is "what is better for X," the question is how to understand the "for" in "better for X". The relativization "for" is "relative to the interests of". But "better relative to the interests of" does not correspond to the notion of "self-interested" as usually conceived. The pains and pleasures of an individual are not usually all that an individual seeks. The interests a person may have can include things like world peace or the success of the Red Sox. Without a reductive tie to given sensations, self-interest is not separate from interests in very many things intuitively separate from the self.

There is more to people's preferences than just partiality to oneself, narrowly conceived. Not only do people very often take their interests in their own welfare more thoroughly to heart than they take the interests of others, but their interests include interests in the welfare of a select group of others.[31] Parents have special bonds to children, friends to friends, and many Americans to other Americans.[32] Such special bonds that A has to B include the interests of B, (where B can be other people or groups), in the interests of A.

There is no sharp line between the prudential interests of a father and the interests of his child. The father's happiness and well-being is at least as closely tied to the daughter's sensations and well-being as it is to any of the sensations that originate in his own internal reports of pleasure and pain. The line between self-interest and the interests of others is not sharp in any individual other than a sociopath.

The identification of interests that flow from human relationships are part of a normal life and part of a good life. If every attachment were determined according to the merits of the case, so that relationships were contingent on the continuing merit of the participants, life would be lonely indeed. More or less unconditional love is important. An interpretation of a person's behavior that attributed no such attachments would require substantial background, essentially interpreting a pathology.

Every culture's "ethical theory" makes some relationships reasons for special concern. Interests in others' interests, then, are partial identifications of the interests that motivate an individual with the interests of others. They are also a kind of deconstruction of the primacy and clarity of the “purely personal” preferences that have bedeviled ethical theories since the beginning. People routinely care about a lot more than their own pleasures and pains, crudely speaking. Partialities can be evaluated as rational or irrational, and we do so routinely in interpretation. A partiality can be excessive, as some allegiances to states and sports teams often are. Degrees of partiality are also subject to reasoned evaluation. A man who is just as attached to the Red Sox as to his children is being unreasonably loyal to the Red Sox. The basis for such judgments is that there is no reason for the partiality. Partiality can be inadequate. A lack of favoritism towards one’s children, for example, calls for explanation by pointing out that the mother is den mother of the group of Cub Scouts of whom her child is a member. So interests in the welfare of others can be rational or irrational.

The maximal scope of concern is concern for the interests of every entity that has interests. Is this maximal partiality rational? If Nagel and Kant are right, it is. If we are motivated by partialities that it is reasonable to have, then this one, “the moral point of view,” is rational, and reasonably ought to motivate any agent with interests.[33] The question is why it motivates so rarely, and so much in disproportion to judgments of “what is better for me and mine.”

b) Why we make this mis-judgment

Given that there is no well-defined subject matter for self-interested desires, and given that people in fact take into account the desires of at least a subset of others, the disposition to focus on limited perceptions of the good must be akin to limited perceptions of the true. That is, just as many people over-rate their own experience in judging what is true, so many people over-rate their own perceptions of the Good in judging what is better than what.

How is this compatible with their valuations being mostly correct? Here is an analogy: Most perceptual judgments are correct; error tends to be in the inferences drawn from those judgments.

In the same way, most judgments of what is better than what are correct, but the inferences are mistaken.

The fundamental “selfish” error is to judge that A and B is better than not-A and not-B because A is good, where A is an involuntary judgment that something is good. That is, while it is correct that A is good, and better than not A, it may not be that A and B is worse than not-A. In effect, people over-rate the goods that are immediate to them or that are experientially connected to goods that are immediate. In the case of immediate judgments of value, involuntary utterances “This is good” “That is bad” carry undue weight in calculating the value of compounds of which this and that are components.

Another way of putting the same position is that agents

“A is better than that B for C”[34] can be treated as a perspectival notion, that is, relative to C’s local valuational beliefs. The analogy with points of view in perception should be apparent. Consider a pyramid with a square base. The pyramid from my angle is square; from another angle it is triangular.[35] So we can coherently take preferences to be beliefs about what is objectively better, and reasonably take “narrow” views on what is better to be due to failures of reason.

If Davidson is right that we mostly agree about the Good as well as the True,

c) Davidsonian Ethics

c1) Davidsonian ethics need not be utilitarian

What is better, all things considered, might or might not be what is better for everybody, and might or might not be a summation of everyone’s preferences. While an identification of what is better, all things considered, with what is better for everyone is a candidate Davidsonian ethical theory, it is not the only theory that could arise from a Davidsonian approach. It could be that what is really better requires something more. Roughly, the relation between what is objectively better and the preferences of agents is analogous to the relation between the truths and the summation of beliefs of agents. While the beliefs have to be largely true, truth is not reducible to consensus. In the same way, while most preferences have to correspond largely to what is really better, the Good is not necessarily the maximal satisfaction of preferences.

C2) Partiality and relativism

Partiality is perhaps the deepest puzzle about morality and its rationality and objectivity. Call a system of partialities a list of partialities that a person ought to have and can have, with a ranking of which partiality should trump which. One troubling fact about partialities is that different cultures have different systems. While (almost) everyone takes certain partialities to be normal and reasonable, there seems to be little way to argue for one system of partialities over another. Partialities, by their very nature, are relative in the sense that they provide reasons for preferential treatment of your son for you but not for me, but not based on any intrinsic properties of your son. One source of irresolvable differences among cultures is the differences in the partialities that are normal in the culture. What, for instance, should be the relation between one's partiality towards one's parents and one's partiality towards one's children? Which should be greater?

A Davidsonian can take one of four positions about partiality:

1) Kant is right, and all partialities are irrational, and therefore morally mistaken. In effect this says that the moral point of view requires not only taking everyone's interests into account, but taking everyone's interests into account equally. Perhaps something of parent-child relationships could be salvaged on utilitarian grounds, perhaps not.

2) One system of partialities, with their rankings of relative importance, is objectively correct. So, given that an American, Fred has promised both his daughter and his father to spend the afternoon with them, and he cannot spend the afternoon with both, there will be an objective answer as to which partiality takes precedence. As observed above, cultures differ in the partialities they regard as normal or obligatory. This position would be that, wherever that situation occurs, like-placed individuals should do the same thing.

3) A system of partialities is correct only relative to a culture. This option would be a form of relativism that allowed that there were some objective truths about what was better than what, but a whole range of states of affairs that were only good or bad relative to the cultures in which individuals lived. Partialities would be valuations that were essentially subjective, about which there was no right answer as to whether the given state of affairs, say that Fred chooses to spend the afternoon with his father, was good or not. It is hard to see how this could be compatible with preferences being judgments that a state of affairs is good.

4) A system of partialities is objectively right for a culture. Given that Fred is an American, an application of partiality may be objectively right even though for a Chinese in precisely the same situation, doing what Fred did would be objectively wrong. This option takes into account that actors are situated in cultures.

This fourth position, to which I am most drawn, is not a relativism. Except in rare cases in which it is indeterminate in which culture an individual inhabits, there is an objective answer to what action is good.

All four positions allow that there is a moral point of view in which every being that has interests is taken into account, and that one should take that view. The options that allow partiality to be rational, however, allow that people ought not to, or that it is morally acceptable not to, treat everyone's interests as of equal weight.

IX The Ethics of Cultures

There are several grounds that should lead us to suspect that there will be irresolvable differences between distinct cultures. Such irresolvable differences are compatible with a great deal of agreement on ethical matters. Some such differences may be cases of one culture getting something wrong, others may be indeterminacies, and others may be case of option 4) above. Some of the sources of irresolvable disagreement are the following:

1) Difference in systems of partiality is one reason that there can be profound irresolvable differences between cultures, as discussed above.

2) Even though two cultures agree that, for instance piety and courage are virtues, they may disagree on their relative ranking. In such disagreement, one may be right, or the question may be indeterminate.

3) Ethical mistakes may become imbedded in cultural practices. If ethics is the science of what is really good, and of what is really better, then ethics is very difficult in application to many concrete cases in which many factors are relevant.

The familiar analogy might be with weather-prediction. Some weather predications are correct and easy. If dark clouds are to the West, it is likely to rain. If it's January in Connecticut, it will be less than 90 degrees Fahrenheit. But determining in detail whether and how much it will snow in Storrs is beyond even dedicated and sophisticated specialists. However, what the weather in detail will actually be in Storrs on a given day is almost certainly in principle predictable.

In the same way, if something like the above account is right, very many choices about what ought to be done are likewise decidable in principle but practically very difficult to determine. The difference in the case of ethical questions is that questions in ethics demand answers in a way that questions about the weather do not. Since we have to do something, we decide.[36] Since we have some rules of thumb, and some intuitions that must be right, we use those rules and intuitions. Much of the time, just as a farmer's weather predictions are, we are right. Another difference is that, whereas everyone can determine when the weather-prediction is wrong, no such clear-cut evidence is forthcoming when the ethical theory embodied in a culture's practices yields wrong results.

The necessity to have answers to ethical questions such that many people can agree on the answers generates the "ethical theories" that are embodied in cultures and which are part of the socialization of members of a culture. When the answers generated by the necessity to have an answer are mistaken, entire cultures can be wrong.

What should we expect when two different cultures have an ethical disagreement on a kind of hard case? No matter how sincere and rational the negotiators are, there is no general algorithm for determining what is the better course of action. Each culture will perforce be using its own conception of the rational. That does not mean that there is no right answer. Sometimes both are wrong, sometimes one is right, but some cases may just be indeterminate.

---

[1] Critics will note that this begs the question in favor of "maximizing agreement," since the presupposition is that our notion of rationality is roughly right.

[2] For instance: Aliens have landed on campus and are choosing whom to abduct for their experiments. They wish to take only those without already existing pain, so as to have an accurate baseline. In this situation, seeking pain would be an explanation.

[3] From this thesis alone, some degree of agreement on ethical matters would seem to follow. Other things being equal, a value one individual holds is a value the individual being interpreted holds as

---

well. Some agreement on values, conceptions of the good, follows from another entity being a rational agent.

[4] Rather than try to give a general proof of this, I will give an example. Consider a person who is tying his shoe. Given an overwhelming desire that a comet crash into Connecticut and an utter lack of concern about shoe-wearing safety, this could reflect the belief that tying this shoe now will bring that about.

[5] In Davidson, “How is weakness of the will possible?”, *Essays on Actions and Events*, Oxford 1980.

[6] An adequate theory of logical form will contain an account of “if”-clauses that covers all their uses. William Lycan’s *Real Conditionals* (Oxford 2000) is an introduction to some of the difficulties.

[7] This point about conditional probability is from Hempel’s “Inductive Inconsistencies,” *Synthese* 12 (1960) pp. 439-469.

[8] “Is probable” is related to “the probability of A, given B is 0.5” in something like the way that “tall” is related to “Is 2 meters in height”. “Is probable” and “probably,” the adverbial form, are attributives on a dimension of “the degree of support the `antecedent’ gives to the ‘consequent.’” “If A, then probably B” means, roughly, “the probability of B given A is pretty high”.

[9] The theory of “prima facie obligation,” which is designed to accommodate this phenomenon while retaining “principles” in the traditional sense, is a much less satisfying account, akin to a theory of probability as “prima facie necessity.”.

[10] Other relativizations may be contextual. After a glance at the sky, I may say “It will probably rain.”

---

[11] An important question here is whether the “ought” is relative to all the information that’s available or to all the truths that are relevant. If, unbeknownst to all, a medium-sized meteor is in fact headed for Cavey’s, and meteor-impacts greatly diminish dining pleasures, then what is reasonable to do given all available information may differ from what it is reasonable to do objectively.

[12] A successful communicator will be aware of contexts and expectations of the audience, so that the intention is made apparent. If the context is ambiguous, a phrase like, “morally” may be used.

[13] Analogous relativization occurs with antecedents in the “logical” “ought”. “If you believe it’s raining and you believe it’s cold, you ought to believe it’s cold and rainy.” Nothing recommends the belief apart from the antecedent, combined with the limitation of considerations at play to the “purely logical”.

In the case of the “logical ought” other considerations that over-ride are even more apparent. If the evidence is very much against rain and against cold, then a person who believes it is raining and believes it is cold should abandon both beliefs rather than believe their conjunction.

[14] The notion of “self-interested” will turn out to be poorly-defined, as we will see below.

[15] What I have called the “logical” “ought,” as in “If you believe A and you believe B, then you ought to believe the conjunction of A and B” likewise has a restricted background. The restriction in this case is that only the logical competence of the agent is under consideration. Thus, the sentence is compatible with there being some terrible consequence of believing the conjunction of A and B. As a general “ought” principle, it also ignores the possibility that A is the contradictory of B.

---

[16] A Davidsonian will need a theory of propositions that will fit the roles of “proposition”. The details of how one does this by applications of Davidson’s “On Saying That” (in *Inquiries into Truth and Interpretation*, Oxford UP, 1980, pp. 93-108) are not clear.

[17] Exactly what the form of conditionals in English is in any case is a matter of much controversy. See William Lycan’s *Real Conditionals* (Oxford, 2001) for a discussion of the many views current and a view of his own.

[18] See Alan Hajek’s “What Conditional Probability Could Not Be,” Nous 2003.

[19] In radical interpretation, for instance, having a “theory” of another’s speech is essentially to have an account of what alternative utterances *would* mean. That is, counterfactuals are admitted at the very core of the theory. Since an account of what a person means on an occasion of utterance is central to the Davidsonian account of meaning, there is no hope for a non-circular reduction of the modal to the linguistic, since the linguistic presupposes the modal.

[20] For instance, the “senses” of “can’t” in which I can’t be at the meeting (because I have a dentist’s appointment), in which I can’t get to China in one billionth of a second, and in which I can’t find a ratio of integers whose square is two. I am supposing that the modal “operators” can ultimately be treated as predicates of “things said” or sets of “things said.”

[21] For “can’t”, the larger the set of “principles,” the larger the class of things said that “”can’t” will apply to. So, in the case of the meeting that I can’t attend because of my dentist appointment, the “principles” will be guidelines about the relative importance of meetings and dentist appointments. Since logical consequence is the device by which all the considerations are brought to bear on the candidates for what can’t be done, “logical impossibility” is the limiting case in which the “can’t” is relative to no considerations.

---

[22] See Twersky, Amos and Kanhneman, Daniel, “Judgment under Uncertainty: Heauristics and Biases”, Science, 185 (1974) pp.1124-1131.

[23] There is something called “obligation,” but that has to do with contracts and promises. The long-standing tradition of identifying morality with obedience to law, specifically the covenants with a divine being, have misled Western moral thought.

[24] Nagel, Thomas, *The Possibility of Altruism*, Princeton UP, 1970.

[25] Given that induction is not reducible to an algorithm, it may also be that inference-patterns can differ in subtle ways that are not part of “content” nor yet part of “structure” in any useful sense. A difference in dispositions of “judgment calls” would be an example.

[26] A rational agent with a very bizarre desire is being irrational, just as a rational agent who holds a contradiction to be true is being irrational. Although for a Quinean-Davidsonian there is no essence to rationality, there are innumerable clear cases.

[27] The good for an individual would be good for an individual with the particular internal sense-equipment that he happens to have. This kind of relativism does not threaten moral objectivity, since there is an objective moral truth about individuals with certain characteristics.

[28] Preferences are strictly expressed in “that”-clauses, so that “A is better than B” would be some appropriate “that F is better than that G”. Little is lost in the brevity of using the noun-phrase version “A is better than B”.

[29] On an account on which there is no given, sensory beliefs, such as my belief that there is a screen in front of me, can be construed as utterances that occur involuntarily, and have to be interpreted. In adapted Quinean terms, such beliefs, by a kind of self-interpretation, have to be incorporated into our web of belief and desire.

[30] It is worth noting that this is essentially Aristotle’s account in the Nichomachean Ethics.

---

[31] Typically, such concern with interests of others is differential, and yields partiality or favoritism. The vexed issue of whether discrimination in favor of one's children is rational, or more generally, whether a concern for the interests of all other entities with interests rationally leads to *equal* concern with the interests of such entities we will ignore. For the moment, I would like to establish that it is rational to have *some* concern for the interests of anything that has interests.

[32] Among the features that are maximized in interpretation, as noted above, are emotional responses and emotional attachments. For a Davidsonian, emotions are subject to evaluation by reason, and therefore part of the content of rationality, since we maximize appropriateness of emotional response in interpretation, just as we maximize appropriateness of distaste or positive preference in the case of sensations. A person who enters his burning house to rescue his Tupperware may be evaluated as having behaved irrationally. So Kant's move of leaving emotion out of morality is not available.

[33] I should re-emphasize that this is a very limited result, since there is no argument that concern with others should be equal. It is implausible that a mother rationally should care about all children to the degree that she cares about her own, for instance. And this implausibility is built into interpretation.

So, a large question about genuine conflicts of interests, when you want something for your child that I want for my child, are being ignored.

[34] "Good," "better than," and similar words have many complexities. "Good pole-vaulter" is not a conjunction, since a good pole-vaulter who is also a shot-putter need not be a good shot-putter. A theory of "good" and "better than" also needs to take into account "good at" and "good as" and well as two ways of using "good for". The theory that most attracts me is that of Richard Larsen,

---

begun in his essay, “Olga is a beautiful dancer”. For purposes of the present essay, I will suppose that some theory of “good” and “better than” is in place, since little hangs on what theory is applied.

[35] I may make a mistake about what aspect the pyramid presents from my angle, by careless observation. In the same way, C’s interests are not just C’s desires. C’s interests are his objective goods, about which he can be mistaken, if he is misinformed about what the objective situation is. There is a difference between what C thinks is better for him and what really is. A correct perspective on the good requires correct information about the true.

[36] A moral skeptic may claim that the better analogy is with astrology, where the King demands an astrological forecast, since the stars must be presaging something. The skeptic will say that just as the right response there is to deny that the stars control destiny, so here the right response is that there is no objectively better in general.