Politics, Principles and Pluralism:

On why liberalism must be inconsistent if correct.

by

Kurt Holukoff

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Philosophy

Waterloo, Ontario, Canada, 2014

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In this dissertation, the author argues that constructivist foundations of political liberalism require a rarely recognized sort of pluralism—not only the familiar pluralism between ideas about how we ought to live that are the stock in trade of standard accounts of liberalism, but a pluralism about political foundations as well. The author argues that making sense of this requires revision to the way we sometimes understand key concepts (such as obligation), and develops an inconsistency-tolerant, pluralism friendly deontic logic for this purpose.

A pluralist friendly obligation is argued to be one that represents moral and political principles in contrastive terms (analogous to contrastive explanation from Bas Van Fraassen), in virtue of the need to order acting upon prescriptions. The author develops a class of mathematical objects *choices* to model answers to why we should choose one policy over alternatives.

Constructivist foundations also turn out to be *prima facie* pluralist foundations, in virtue of the nature of the norms guiding abstraction. This leads to a proof that, in a weakest base logic, legitimate moral or political codes in a pluralist context must reference each other. Upon explicating the distinction between perspectives that could consider unrealizable plans and perspectives that are themselves unrealizable, the author proves that in our world liberalism is itself an unrealizable plan. These results clearly illuminate what is at stake when justifying foundations for a liberal state.

Acknowledgements

Table of Contents

Politics, Principles and Pluralism:

On why liberalism must be inconsistent if correct.

I. Introduction: Justification and Pluralism

  Justifying a liberal approach to political policy, especially fundamental policy, begins with taking the deep divisions and disagreements in a society seriously. In other words, *liberalism* is based upon balancing and prioritizing the moral and life concerns of the various members of a pluralist society. In this thesis I shall argue that there is important philosophical work that needs to be done before we can get to this fundamental balancing act, i.e., deciding upon some basis for justifying political liberalism; there are good reasons to think that we need to sort out what some of the concepts we routinely use in political and moral discussions can mean in a pluralist context. To be clear, I am proposing that many of the concepts we take for granted when theorizing liberalism may be unfit in a pluralist context. My hope is that if we can identify exactly how and why this mismatch occurs, substantial progress will not be far behind. Let me begin with some motivating examples.

I.i Ought and Can

  Is it possible that 'ought' implies 'can' is a platitude? If not, under what conditions does 'ought' imply 'can'? If a pluralism regarding correct or acceptable ethical theories were plausible, such conditions would make the connection between what we ought to do and what we can do far more complicated than may be surmised. The problem is, at least at first glance, if one hews to a meta-ethics allowing more than one

1

correct but conflicting moral theory, one ends up committed to the claim that 'ought' does not imply 'can' except in those [rare] occasions where there is only one such correct theory that is actually applicable.

A brief sketch of this argument: suppose we have two equally correct, but conflicting theories of what we ought to do. Since they conflict, it must be that they prescribe actions that turn out to be incompatible, i.e., acting on one obligation precludes being able to do the other. So we have a case where we ought to do two actions, but can only do one. Thus, we ought to do something but can't do it. Obviously, this sketch calls for a more detailed discussion, which can be found in chapter two.

I.ii Consistencies and Equilibrium

As is well known, reflective equilibrium prescribes *mutual* adjustment between general principles and concrete judgments, with the aim of achieving not only a sort of consistency between abstract principles and exemplary cases, but also a best version of such a balance. The idea here turns on two subordinate claims. First, that there are better and worse ways of reaching this balance. Second, that we identify a best such balance with a justification for the constitutive principles and judgments regarding the exemplary cases. In other words, it's insufficient for us to merely obtain consistency between principles and intuitions; different approaches to reaching a balance, i.e., a consistent package of principles and intuitions, need to be taken into account and compared. If we

can't do better[1] than a given balance of principles and intuitions, those principles and intuitions are correct (or justified).

Arguably, pluralism doesn't play nice with reflective equilibrium either. Here is a sketch of why. Pluralism entails irreducible disagreement between equally correct theories—see chapter one for details. Thus we are guaranteed at least some inconsistency between the principles we expect to take as correct in some cases, concrete or otherwise. But the previous explication of reflective equilibrium made (seemingly essential) reference to consistency. Thus, if we use reflective equilibrium to justify theories/concrete judgments in a subject matter where we think a plurality of theories may hold, it seems we are doomed to failure.[2] If you can't get a consistent group of principles and intuitions, it seems to follow that instead of a plurality of correct theories, we instead have none.

Here is another way to approach this idea. Let's suppose that we have a reflective equilibrium set up that looks like we have multiple best balances of principles and intuitions. If this could make sense, each such principle and intuition could be justified as correct according to proponents of reflective equilibrium. On this supposition, there will be multiple best but conflicting balances between various principles and intuitions: maximals rather than a maximum. Unfortunately for those wanting to treat reflective

---

[1] And of course, I will remain silent on what makes one version better than another. Readers may fill in the blank as they see fit, nothing in what follows should hang in the balance.

[2] This is how I understand Nelson Goodman's later argument that there must be a plurality of actual worlds [Goodman 1984]. Since he argues that a version, i.e., a philosophical description, of a world must be consistent, it must mean that there are many such worlds since there are many such correct descriptions. These theoretical descriptions just happen to be mutually inconsistent. Perhaps, this is why he claims to be a relativist.

equilibrium as a justificatory method, since these local maximals will conflict with each other, the principles they espouse can't be consistent with each other. And thus, reflective equilibrium calls for further adjustment (and perpetually so). In other words, a reflective equilibrium approach can't seemingly allow us to claim that more than one, but conflicting, theory is correct regarding the same issue in the first place (i.e., if consistency must be a feature of whatever sets of principles and intuitions from which we then look for a best such version). Thus, there are no consistent solutions possible between all the principles (and intuitions) we think should be correct, given pluralism.

While obviously not conclusive (there may be some way to relax the requirement of consistency in pluralist contexts[3]), the above sketch of an argument is intended instead to be only suggestive. As such, I do not defend this argument in detail in what follows. Rather, the point here is that meta-ethical pluralism should not be accepted blithely, i.e., just because we think that there is theoretical value to allowing multiple but conflicting theories or values/principles. This said, I propose that we should be at least suspicious regarding the coherence of pluralism in meta-ethics and thus, there is a pressing need to understand just what we are committed to when intuitions incline us to theorize liberalism in pluralist terms.

II. Liberalism is a Pluralism

Of course, if pluralism seems incompatible with what we think about core ideas, one might think *so much the worse for pluralism in meta-ethics*. And philosophically, if

---

[3] Perhaps use some version of pareto optimality instead of best equilibrium or even reject pluralism by compartmentalizing the different theories so they are talking about different subjects, i.e., treat each local maximal as its own type.

that is the way you see the conceptual costs and benefits tilt, so be it. After all, if pluralism in meta-ethics requires generally jettisoning 'ought' implies 'can' and the use of a basic philosophical tool like reflective equilibrium, we better look for something else besides a pluralist hypothesis. However, at least for those inclined to liberalism, this is not a viable response—pluralism in any meta-ethics compatible with liberalism runs deeper than may have first been thought, a claim I now turn to argue.

In the foundations of political liberalism we find a sort of pluralism that refuses to go away. And trying to formulate the foundations of liberalism without taking this pluralism seriously has led to some notable, even if fruitful, failures. Before he died, John Rawls argued that merely taking pluralism about what is good in life seriously is not enough [Rawls 2005]. In order to account for all that a liberal theorist should, one must accept a plurality of what he called Public Reasons; in particular his principles of justice should be just one type (and perhaps a minor one at that) of Public Reason among many (pg. 450). To avoid the use of his particular jargon, we can say that Rawls recognized a plurality of general principles of political justice. But why take Rawls' advice? The answer is that formulating the foundations of liberalism requires taking reasonable disagreement seriously, at whatever level of analysis one may find it. Here's how.

II.i Liberals Disagreeing and Obligation

The principles of political liberalism are supposed to lay out how we should live together while still disagreeing, and even fundamentally so, about what is good in life. Such disagreement about what is good in life may be thought of as pluralism about

private morality, as opposed to public morality.[4] At this level, correct liberal principles should then play a dual role, providing both a moral justification for a particular approach to the problem of how to resolve tensions arising from differing private moralities, i.e., pluralism about the good life, and also a providing a degree of practical guidance about how to do so.

But, as we will see, there are good reasons to think liberalism is itself a pluralism, not just a single story about how to deal with *private* morality pluralism. What does it mean that liberalism is a pluralism? Since it is the *ambition* of political liberalism[5] to set up how we can still cooperate *qua* society, given the disagreement found in pluralist society, the liberal theorist's job is to provide a correct account about how to appropriately enforce such cooperation. Of course, different theorists propose different, incompatible answers to both aspects of this question: the terms of cooperation and its enforcement or implementation. Liberalism is a pluralism if these different theorists are equally correct about how we should cooperate (and enforce it).

The launching point for this project is then to ask what seems an obvious question, framed this way: why can't people reasonably disagree about how we should cooperate

---

[4] While there may be different ways of spelling out this (often fraught) distinction, for our purposes what we want to be able to distinguish is cases where the principles or values involved regulate the use of resources (time, money, power etc.) provided in trust by others, i.e., a public case, from cases where such issues arise at best, if at all, derivatively, i.e., a private case. The question of what ought I do in public cases must appeal to norms involving the reasonable expectations of those who have entrusted the relevant resources, while in the private case, one is freed from having to account for the expectations of others as they see them. I.e., in the private case one may account for others according to one's reasonable moral theory or considered judgment. But in the public case, one ought to account for the relevant expectations on their own terms, or more precisely a *negotiated* version of their terms.
[5] Articulating this notion in terms of an *ambition* is due to a suggestion from Mathieu Doucet in conversation.

qua liberal society, and not just about what is good in life? To be more direct, it may have seemed safe to assume that it is counter-productive, if not impossible, to theorize the liberal state by using a variety of correct, complete, but disagreeing accounts of fairness, or justice. (Or whatever else one takes to be the fundamental notions underpinning the rules of cooperation *qua* society for a suitably *liberal* state.) After all, liberalism is supposed to resolve how to get along while disagreeing; and disagreeing about this might be tantamount to the failure of the liberal project itself. But this 'no-disagreement' assumption for the foundation of liberalism I shall argue is untenable. In fact, in chapter one I shall argue that this is an instance of a more general philosophical lesson: constructive foundations in general (i.e., in any subject matter) are also, *prima facie*, pluralist foundations.

Put in these terms, it may be thought that this line of argumentation compounds the problems of liberalism. That is, granted we take seriously the suggestion that pluralism is of dubious coherence (for example on grounds like its seeming incompatibility with certain truisms), the last thing we should want is to theorize liberalism in terms of pluralism. But it seems to me that we need to take the question of 'why can't we reasonably disagree about how to cooperate' as, at least *prima facie*, a serious one. Consider, for instance, some debates about the rules of cooperation as they actually occur in contemporary liberal societies. Let us briefly discuss the debates regarding the best system for electing members of legislatures.

II.ii Unfair/Fair Elections

Readers are probably somewhat familiar with various options for what we may call *electoral formulas*—first past the post, proportional representation, automatic run-off and others, some of which have several variants. What I want to note here is that each electoral formula has a plausible case to be made for it on grounds of fairness and justice. Moreover, each is also susceptible to criticism on grounds of fairness and justice. For ease of presentation, consider just a few of the various negative criteria for a fair and just electoral formula or system:

It is *prima facie* unfair or unjust:

a) to ignore/abstract away anyone's vote in deciding which candidate wins a seat

b) to ask citizens to vote without suitable political debate (and this requires local debates)

c) to not let citizens vote for a local representative if they so choose

d) to allow the system to abstract away which party wins the most 'first choice' votes in deciding which candidate wins the seat

e) for a system to allow seat distribution to be unrepresentative of voters' preferences, i.e., a suitable sample of seats should be representative of the voting population, suitably construed.

For our purposes, what matters is that it is unclear whether the burden for any relevant criterion is met for any of one of the electoral systems. And I am sure there are many more criteria that may cloud matters further. Here are a couple of examples. With *first past the post* in Canada, for example, generally half of the time you only need little better

than a third of the votes in a riding to win the seat.[6] So while FPTP allows and encourages local debates, with local representatives and ensures at the riding level the candidate who wins the most 'first choice' votes wins the seat, half of the time two thirds of votes are ignored and can allow for parties to win a majority of seats while receiving less total votes than another party. On the other hand, Automatic Run-off systems can allow (while in practice rarely do so) that a candidate who wins the most first-place votes to not win the seat by treating first-place votes as, in some sense, on par with 2nd or even lower place votes. And of course, there is nothing in either system that straightforwardly encourages or ensures a legislature or executive that is demographically representative of the voting population.[7]

It seems that taking some types of fairness considerations seriously will show any one of these electoral formulas in a favorable light but at the cost of disregarding other types of fairness considerations. The question of what *makes* an electoral system fair seems to have multiple cogent answers. Each sort of consideration from the above example emphasized different characteristics we take to be relevant when deciding whether a voting system is fair. If this characterization of the fairness of electoral formulas is cogent, reasonable disagreement about the right way to cooperate, i.e., in terms of this sort of basic institution, will be inevitable.

---

[6] Accessed June 10 2014
http://www.citizensassembly.gov.on.ca/assets/From%20Votes%20to%20Seats.pdf
[7] Again, the electoral formula or system itself doesn't encourage demographic representation, but if some ridings are dominated by demographic subsections parties have an incentive to use candidates representative of that demographic. But this is happy coincidence, not a virtue of the systems in question.

Of course, I don't pretend that one example, or even a range of similar ones, could establish that this sort of reasonable disagreement exists—at this stage I aim only to be suggestive. What this suggests is that liberalism is caught between needing to take disagreement seriously at all levels and on the other hand maintaining theoretical coherence. I assume that this problem is not merely 'eating cake and having it too'; rather, if we want to articulate how liberalism can itself be a pluralism[8] we need better conceptual tools. This project aims to develop a foundation for the concepts we need to correctly theorize a truly pluralist liberalism.

III. Obligations and Pluralism

To develop a foundation for a thoroughly pluralist liberalism, this thesis is roughly divided into two parts. The first section examines the nature of pluralism and reasons for thinking when and how it could be true, in general and in moral domains specifically. In these sections I also explain how the formal and informal methodologies I have chosen are well suited for analyzing the relevant notions. That is, I show that a pluralist context makes certain demands on how we theorize.

The second part pursues the task of providing a pluralism-compatible account of the central moral concepts it seems to conflict with. If the dubious coherence of pluralism results from a faulty analysis of the concepts, like *obligation* (especially as it relates to

---

[8] One way to motivate research into the various Relevant logics is to take the paradoxes of self-reference (such as the various Liar sentences) as valid with true premises. Articulating liberalism as a solution to private morality pluralism but as a pluralism itself is likely no accident. You can find in chapter four a proof that in pluralist contexts moral codes must refer to each other. However, the exact nature of the self-referencing nature of liberalism remains to be worked out in future work.

the notion of *principles* of justice in a liberal democratic context), an account of *obligation* and its cognates that can make sense in a pluralist context is desirable. To develop one, I defend a contrastive analysis of *obligation* under conditions of pluralism, which allows us to make sense of disagreement about the right way to socially cooperate.

I also assume that the principles of liberal justice are decided upon, not discovered; in this sense, my approach is part of the constructive traditions in liberal foundations and so my approach can be seen as offering an assist to such liberal theorists who have either hinted at or advocated pluralism about liberal foundations. Moreover, as pointed out already, if my arguments in chapter one, i.e., that constructive foundations are *prima facie* pluralist, are correct, my approach is important for any constructivist in the liberal tradition.

To summarize the general point another way, in this project I answer two questions at the heart of this pluralist turn. The first question: How is it possible (or more precisely, what would it cost) to have clear and intelligible notions of *obligation, moral principles* and other cognates if moral or political pluralism is assumed? And supposing a satisfying answer to that question: is there a suitable notion of pluralism that, in fact, could help constructivist liberals develop better a foundation for liberalism?

IV. Method: Formalizing Obligation

While there are various moral concepts the obvious readings of which conflict with pluralism, much of my discussion focuses on obligation. I do so since it is a linchpin

or keystone concept, which once clarified will make the needed adjustments to other notions clear. A simple sketch should suffice for now.

The political rules arising from a particular foundational theory will tell citizens what they *must do, ought to do* or are *forbidden from doing* (e.g., in order to be fair to everyone) and what they are *allowed to do*. Presumably in a liberal state, the final category is large, the first three smaller, as befits a theory with the notion of liberty at its core. But if there are in fact various correct such foundational theories, and they are genuinely *distinct and so disagreeing* theories, presumably there will be disagreement about which category certain actions fall into.

It is tempting to take such inconsistency as motivating a disambiguation rule. That is, to avoid saying that a particular action is both obligatory and not obligatory, or that two incompatible actions are both obligatory, one might defend the view that we mean different things by the notion of *prescription* found in each endorsed moral code. Rather than a plurality of equally correct but conflicting moral codes, we have only different theories with different conceptions of *prescription* and so no disagreement or pluralism after all. I will pursue a different road: if you get the right logic, inconsistency does not indicate the need to disambiguate.

For a few decades now, logicians have pursued formal means of reflecting the intuition that all inconsistencies aren't the same. Some sorts of inconsistency are not disastrous, logically speaking. The best-known sorts of logics that can distinguish between disastrous and deductively interesting inconsistencies are the various Relevant (or Relevance) logics. Of course, Classical and Intuitionistic logics treat all

inconsistencies as explosive, implying everything and nothing. Representing deductions in a pluralist context however requires us to distinguish between when someone uses a particular theory in an inconsistent manner, which should be explosive, logically speaking, and when someone uses inconsistent claims from two different but equally correct theories. The latter sort of inconsistency must be allowed for, i.e., to put constraints on what follows from such inconsistent group of premises.

The use of ready-made tools from formal logic is one reason that there is more symbolism in this thesis than one often finds in works of political philosophy—others have already considered how to revise what follows from inconsistencies. But there are further reasons to take a formal approach. Formalization is well known as the appropriate and go-to tool for investigating conceptual options and constraints for the sorts of concepts that can in some sense be treated as *logical* words, and words like 'ought' have long been regarded this way. A proper formal treatment of these concepts will help us sort out what these concepts can coherently mean in a pluralist setting.

As may be guessed, standard deontic logics are not the right tool for the present job since inconsistency is disastrous in such logics; nonetheless I shall briefly argue this point in chapter two. The positive part of the project, so to speak, and the one the vast majority of this thesis deals with, is that we shall need to develop new tools for thinking about obligations (and cognates) in a pluralist setting.

V. The Difference between the Political and the Moral

I am arguing that justifying liberal, fundamental policy requires understanding what justification would look like under conditions of pluralism. To accomplish this, there are two tracks I take in this project. First, as mentioned, the methodological track uses (or aims to use) a suitable notion of deontic consequence, i.e., a logic that treats 'obligation' as a logical word. The philosophical track is to suppose that, as we are searching for norms that we can use to justify political liberalism, we can use the sort of norms we already use in the moral domain and make suitable adjustments for use in the political sphere. In both cases, the adjustments I will argue for will be due sometimes to considerations regarding the nature of value pluralism and other times in regards to special features found in political claims (or at least allow for such a possibility.)

The worry here is how can I allow for the possibility that political claims about liberalism are not the same as substantive moral claims, i.e., liberalism as a political theory needn't assume any particular meta-ethical view. There's a potential tension between using a formalism that focuses on an ostensibly moral notion, as well as often using moral examples to illustrate the norms I argue we should adapt for use in the foundations of liberalism, and respecting the distinction between substantive moral claims and political claims.

The response to this concern in the methodological arena starts with the idea that while it is also a live debate whether pluralism is true in the ethical realm, it is not an implausible suggestion. My account of issues, like one norm trumping another, spell out the need for representing these types and patterns of conceptual reasoning by positing

14

moral pluralism, and account for the mechanisms of, say, trumping, in that light. If I am right, the liberal political domain is also one in which we find a plurality of normative codes, so it is plausible that analogous formal tools will work in both domains. The idea here is that on a substantive or material level of analysis it's possible to draw carefully circumscribed but fruitful analogies between moral and political claims; on a formal level, such analogies are easier to come by, since structural similarity comes much cheaper than does the material. This is generally one of the great virtues of formal analysis, i.e., that the same structure may be instantiated in different (and sometimes radically different) realms, which allows us to use claims in one domain to articulate a similar claim in a different domain.

On this account, we may see moral and political codes as both species of a more general notion of a normative code, and that to a significant degree the structure found in moral codes under conditions of pluralism generalizes and thus can be applied in the analysis of political codes (under conditions of pluralism.) Part of the justification for this attitude (towards the generalizability of the structure of moral codes) is in fact how apt such structure turns out to be, i.e., for representing political codes under pluralist conditions. In other words, I am asking that the context of discovery serve also, at least somewhat, as a context of justification. And of course, if I am right about this species/genus differentiation, we could do no more to philosophically respect the distinction between political claims and substantial moral claims. And all this while nonetheless treating such claims as structurally similar in the way presupposed by this methodology.

Regarding the other aspect of this concern, when I explain what is philosophically or materially at stake regarding certain distinctions, I will generally use moral examples to articulate a first approximation. And given the methodological track I take, I ought not do otherwise, i.e., I need to first formulate a moral version before I can explain why I need to make the required adjustments to render it properly political in nature. Moreover, there is nothing claimed here *a priori* about the nature of political claims, i.e., that they just are notions adapted from moral contexts. Rather, the claim is that I can use the analysis of notions found in the moral context to develop an analysis for the relevant political claim. Here I am instructed by Feigl [1968]: philosophical analysis is the search for the basis of justification for the relevant class of claims. On this account, justification is giving reasons why you choose, believe, etc. *A* instead of *B*: *A* falls under the relevant norm, say *W*, while *B* does not. And if someone demands justification for the norm *W*, none can be given, or, as I would add, not without specifying alternative norms, say *Y* or *X* and arguing that *W* does a better job than *Y* or *X* in regulating what matters in that discourse.

As Dummett [1980] points out, reporting correct usage, i.e., as falling under the relevant norm, is not philosophy, but mere description (pg. 2.) There is a reason for these norms, the 'why we follow said norm' and this reason is what is properly philosophical. In other words, what we do by following the norm is what is at stake, philosophically. As it turns out, at least in my approach, the philosophical stakes are far more easily drawn when dealing with the toy models and thought experiments common to the moral domain. Again, the claim I am defending is not that there is no distinction between the political and the substantive moral; rather that they aren't so different that we can't make suitable

16

adjustments to norms in one domain to justify claims in the other. And this is possible because, as I will argue in particular cases, why we follow these philosophically significant norms are relevantly similar in both domains.

While this argument may not be conclusive in some *a priori* sense, the point is instead to motivate at least taking this sort of methodology seriously enough to test out its results.

VI. Outline

The dissertation is organized as follows. Chapter One, I introduce some of the basic notions underlying this project. First, I introduce the philosophical issues that I take to motivate the move to pluralism in meta-ethics: Bernard Williams' anti-systematic claims regarding ethics and Amartya Sen's *infeasibility* and *redundancy* objections to constructing ideally just states in liberal theory. I then sketch a brief and informal version of my response to these objections, i.e., to using ideal versions of a state to theorize liberalism. I then look at the concepts of deontic logic and the value of using a logical approach to meta-ethics. Next I examine the connection between action-guiding principles, moral constructivism and pluralism. Here I argue that a constructivist foundation is also *prima facie* a pluralist foundation.

In Chapter Two, I examine the nature of pluralism, both in general and in meta-ethics in particular. I first survey some common and pertinent responses to moral disagreement or dilemmas and argue that dilemmas ought to be taken seriously, not explained away. One way to take dilemmas seriously is to posit that there are multiple correct moral theories that conflict in such cases. The next section is a detailed

17

examination of the nature of theoretical pluralism and how it could apply to regimenting

meta-ethics. The next three sections examine a family of what I suggest are logical and

philosophical consequences of taking moral pluralism seriously: inconsistency of

principles, paraconsistency, and (perhaps) true contradictions. From this discussion, a

clear picture of what taking pluralism seriously involves emerges. We can now begin the

analysis of *obligation* and its cognates under conditions of pluralism.

Chapter Three is concerned with the problem of action-guidance in a pluralist

context. The idea is that, when faced with the sort of disagreement seen as problematic

for action-guidance, the relevant codes may prescribe trumping the original but

conflicting obligation with a coordination mechanism. These mechanisms could allow us

to settle disagreement in some, if not most, cases in a liberal society. There are two

subordinate issues involved in this strategy. First, what is a trumping relation? Second,

what is a fair coordination mechanism?

Given these issues, Chapter Three is largely focused on the philosophical details

of how we should analyze *obligation* such that we can make sense of *contrary to duty*

(CTD hereafter) structures and what I call 'trumping' relations. A trumping relation holds

between two obligations when one obligation overrides the other. A CTD structure is

closely related, and the idea here is if one were to do wrong, not all bets are off so to

speak. After one has done wrong, there are still moral reasons to refrain from other evils.[9]

---

[9] The Gentle Murder Paradox is likely the most famous CTD structure in the literature [Forrester 1984]: Suppose you are going to murder someone (clearly contrary to duty). But surely, one ought to mitigate harm when one can. Thus if you murder someone, you ought to murder them gently. The problem is that if one murders someone gently, one murders him or her. And if you ought to do action *A*, you ought to do what *A* entails. Thus you ought to murder him or her. Since we take

The first section details the problems that result from using propositions as basic parameters, i.e., atoms, in any deontic logic that has trumping relations or CTD structures.

The next section argues that moral principles should be understood as more fine-grained than is the usual practice. To capture this more nuanced conception, I defend a view that moral principles need to be understood *contrastively*. While I claim no commitment to the view as originally developed for scientific explanation by Van Fraassen [1980], I do make use of certain elements as found in his proposal. I argue that what matters for representing moral principles is being able to identify what sort of choices one ought or ought not to so choose. To this end, I propose that choices should be represented by a list of options and that principles should be understood as a claim that such and such option should be chosen, morally speaking, *instead* of the other options. In following the literature surrounding Van Fraassen's project, I call my proposal "*Contrastive Principles*". I show how we need a structure that allows us to specify the contrast class of the action we are morally evaluating in order to represent *choices* and thus *moral principles*.

In Chapter Four, the previously informal concepts of choices and obligations, specifically the resulting ideality ordering and related moral orderings are treated formally. I then introduce a proof-theory to reason about and with the concepts defined by this order theory. I show how we can use the weakest formal language to deduce a criterion for any moral code under conditions of pluralism: that any such code must guarantee that acting upon its prescriptions is compatible with the obligations of rival codes. As may be guessed, this criterion bears a strong resemblance to parts of Rawls'

conditionals to be transitive, we have the paradox: If you murder someone, you ought to murder him or her.

inference from what he calls the fact of *Reasonable pluralism*. But of significant theoretical value, I show that this conclusion holds without using any of his theoretical apparatus or assumptions; the bare meaning of logical terms under conditions of pluralism entails the criterion.

With these tools in place, the formal semantics for the language *RD-choice* is introduced. *RD-choice* is shown to be paraconsistent, relevant, and constructive.

Next I introduce a semantics for a pluralist version of obligation: neighborhood semantics. If a sense of permissibility is treated as interdefinable with obligation, I demonstrate the need for a different sense of permissibility for sub-ideal worlds such as ours. This result can represent the distinction between weak and strong permissibility as found in the deontic logic literature. It is also shown that in the neighborhood semantics Williams' [1981] maxim is right; the relationship between *ought* and *must* is the same relationship as that between *best* and *only*.

I conclude with comments regarding directions for research in solving problems in deontic logic, such as CTD structures and other order problems. I also point to the possibility of deriving radically different types of foundations for liberalism. It seems safe to assume that some notion of impartiality must underwrite the foundation for justifying liberal policy. I propose that another sort of foundation is possible, given the arguments regarding the nature of constructive abstraction and pluralism: the fiduciary perspective. Finally I propose that my account of a plurality of moral and political codes, i.e., the type and nature of interaction between such codes, can provide a useful scaffold for theorizing the legitimacy of the state. In brief, when the political codes of a state are used in moral

criticism to adjust private moral codes and when private moral codes are used in moral

criticism to adjust the political codes of a state, that state is taken to be legitimate.

Chapter One   Foundations

1.0 *Introduction*

In this chapter, I will be looking at starting points. The first issue: Bernard Williams and Amartya Sen may be seen as arguing that taking pluralism seriously is tantamount to rejecting systematic constructivism in ethics. If they are correct, this project (i.e., of using systematic constructivism in a pluralist context) is misguided. I will argue that their criticism leaves open at least one avenue of escape, and that this is the option I will take: allowing for a constructive system to be inconsistent. The first problem I look at is the anti-systematic program of Bernard Williams [1981], perhaps best captured by his slogan "'ought' is related to 'must' as 'best' is related to 'only'" (pg. 125.) The other issue is Amartya Sen's [2009] objections to social contract foundations of liberalism. I show how we can use Sen's complaints to make a more general point regarding the problems faced by any constructivist approach. In this chapter I will briefly and informally sketch out how the sort of pluralist approach I propose will be able to respond to these objections. A more detailed response, in some respects, must wait because the details depend on some of the formal work to be done in later chapters.

The second issue follows up on this discussion of pluralism and Sen's and Williams' criticisms: why use a logical approach to meta-ethics? The idea is that while we may want to use a *pluralist perspective* to defend a foundation for political liberalism, we first need to justify a method for investigating the details of a suitable pluralism. It would be cold comfort to have a doctrine that solves problems at a general level, but turns out to be too problematic in the details. To this end, I argue that the relevant moral terms and concepts can be treated as particular sorts of mathematical objects. Given an

appropriate formal language to reason with and about such objects, we will be well on our way to understanding the ramifications and consequences of a pluralist approach for liberal foundations.

The next issue will concern doubts regarding the epistemic status of pluralism in meta-ethics. It may be hoped that the multiplicity of moral approaches to liberalism is merely an artifact of the state of current scholarship, *contra* Williams or Sen. The hope is that once we have better understood our notions, we could develop one uniquely correct theory that captures everything we thought mattered. However, I argue that the very nature of constructivist foundations, when properly understood, generates a *prima facie* case for a plurality of such foundations. So whether or not one agrees with Williams' and Sen's arguments for taking pluralism seriously, as constructivists we have at least a *prima facie* case for so doing. It would appear that a constructivist approach leads one to allow for inconsistent systems in general and in meta-ethics in particular.

1.1 The Organizing Issues: *The distinction between 'Must' and 'Ought'*

To understand why a pluralist turn looks to hold promise for solving issues in the foundations of political liberalism, I will first look at why some have argued that pluralism is instead deeply problematic for systematic constructivist foundations. I begin with Bernard Williams and his criticism of systematic approaches in general.

As in other fields, foundational theories tend to include answers to questions regarding the nature of characteristic concepts (for example, what is the nature of *obligation*). Also, it is common that foundational theories should vindicate practices as

found in its target area.[10] And when we think of *foundations*, thoughts of a systematic

theory are often not far behind. However, in his book *Moral Luck*, and throughout much

of his *oeuvre*, Bernard Williams used features found in concrete moral practices to

support his contention that no systematic foundation was possible for moral theory. I

propose that such concrete moral practices instead suggest that no *self-consistent*

systematic foundation is possible. If I am correct about the nature of these starting points,

my conjecture is that Williams' reasoning should lead us not to reject a systematic

approach *simpliciter*, rather only to reject self-consistent systems. However, while I will

gladly borrow the concrete starting points he describes, the details of Williams'

arguments are beyond the scope of this project. All I need from this discussion is the

conclusion that at least some systems, for example a suitable pluralist foundation, are

permitted, given the relevant concrete starting points.

A major starting point for Williams' [2012] argument was the fact that the meta-

ethical perspectives of his day failed to take seriously the fact that we find irreducibly

conflicting responsibilities throughout lived moral experiences [2012]. His anti-

systematic perspective relied, in part, on a distinction between *ought* and *must* observed

in such concrete instances of conflict. He formulated this distinction as follows: that

"*obligation* is related to *must* as *best* is related to *only*." (Hereafter I may refer to this as

the *obligation as best options* claim.)

The problem the *obligations as best options* claim raises for a systematic

foundations is that it seems if the system involves a unique moral theory, then what it

---

10 Which, in *sotto voce*, often really means "gloss over any revisions."

finds to be the best action, decision, choice etc. for some moral situation is also the only such best option. Thus, there could be no distinction between *ought* and *must*.[11]

What I take from this distinction: if a systematic approach to action-guiding principles is self-consistent, in that for any one particular moral issue there is at most one[12] applicable principle to guide action, the distinction between *ought* and *must* is obscured or even conflated. In other words, when what we ought to do is the only thing we should do (and *vice versa*), there is no real difference between what we must and what we ought to do. Perhaps this suggests that, on this account, the distinction between *ought* and *must* is merely grammatical or a matter of emphasis at best.

Given that we want this distinction to be more meaningful, it seems that correctly abstracting principles from the concrete level of lived moral experiences often requires positing a multiplicity of ways to best live up to our responsibilities. If we want to take into account all relevant moral concerns i.e., found in more concrete terms, then correctly balancing of the responsibilities of different moral roles may take many different guises. Looking at this issue in this way suggests a theoretic possibility with important implications. There is some plausibility in the notion that a theory could have a principle that in some cases finds two states of affairs to have the same amount of good and bad for different reasons. For example, if one is to maximize the amount of happiness within some group, there may be many different ways to calibrate this hedonic calculus. In other

---

[11] True, there is another possible sense of 'must' in ethics: the *must* of a minimum requirement. However, a similar sort of issue arises once we countenance the possibility of minimals instead of a unique minimum.

[12] Technically, the set of evaluations, given the standard notion of a systematic foundation, is a linear chain, i.e., it has a non-branching structure: any option it evaluates is better or worse than any other element it evaluates.

words there may be no unique distribution for the maximum amount of happiness throughout this group.[13]

However, it is not clear whether the lack of distributive principles constitutes a principled distinction between *ought* and *must*. Failing to guide distribution may merely amount to a lacuna in the action guiding power of the theory after all. Nonetheless, we can see from this case that it might be possible to distinguish a choice with multiple best options from there being only one such best option. But the question now is: what would it take for there to be a principled connection between many/one best option(s) and the distinction between *ought/must*?

Perhaps we should look to a more fine-grained decision procedure for a basis for the distinction between *ought* and must. A decision procedure is a principled basis after all. If, as is commonly thought, a constructivist approach is supposed to specify a procedure for identifying the relevant principle for deciding a case, that same procedure should always identify the same principle in all morally similar cases. This consistency over similar cases is a standard, and some might think bedrock, criterion for a cogent decision procedure. Put this way, it seems that there shouldn't be 'tied for best' in constructive procedures.[14] That is, it seems impossible that a consistent procedure could give us different results from the same starting points. The natural option is to then to allow for multiple procedures for deciding what is best.

---

[13] To be clear, while the different distributions may conflict, there is no conflict between moral principles on this account.

[14] Notwithstanding Rawls' contention that the nature of different groups does not matter in distributive justice as long as the proper pattern of distribution holds (i.e., that the worst off are the best off they could be) [Rawls 2005]; however, as will be argued in the next chapter, this notion actually obscures a requirement in his early theory, that is a theory of distributive justice is on the hook for identifying what should count as a fair coordination mechanism among the fair or just distributions.

The lesson, then, from the lack-of-distributive-principles case is that we're looking for a constructive system that, in some cases, finds no moral distinction between multiple options, each relevant to the one and same decision. That is, given at least two such options there should be a relevant and correct decision procedure, which says each is the best thing for us to do. Thus, we need a constructive system with multiple, equally correct but conflicting decision procedures such that there could be no moral reason to prefer one decision procedure over another, at least in some non-trivial cases. Such a possibility would allow us to differentiate between cases where there is only one best option and cases where there are multiple best options; thus we could have a principled *ought/must* distinction.

Of course, if I am correct and theories must be able to be thought of as consistent sets, there must be a way to conceive of a systematic foundation as not a theory. To do so, I will distinguish between a systematic approach and a theory, but this must wait until I explain why and how I will use formal methodology. Before that discussion, let me first introduce Sen's objections to constructing ideally just states as a basis for justifying liberalism. After all, Sen asks a similar question to the one I raised in the introduction: why can't people reasonably disagree about what is unjust?

1.2 The Organizing Issues: *the Infeasibility and Redundancy Problems*

Liberal theory *qua* political doctrine, i.e. over and above the meta-ethical issues regarding pluralism raised in the last section, faces at least two foundational problems, to which I contend, theoretical pluralism instead provides a viable solution.

Let me introduce some background to the issues. It is useful to categorize, broadly speaking, two main strands in the kinds of answers political philosophers give to the question of what policy should guide us in the selection of public policy [Sen 2009: pgs. 5-12]. The received view may be thought of as Regulatory Ideal Theory (RIT) methodologies. The basic idea behind RITs is that from among the realistic alternatives for the basic constitution of a state, we ought to choose the one that most closely approximates the ideal. In this project I am focusing on a subset of RITs, that is, those that use a constructivist framework. To be clear, constructivist RITs specify criteria for acceptability of principles under ideal conditions.

The main challengers to RIT may be called Comparative Approaches (CA). Such approaches argue that working from a neutral perspective, which will vary from theory to theory, the various options could be cardinally ranked according to different standards for 'better than' (or left argued for in informal terms). As described, these are very broad categories. Plato, Kant and Rousseau are some of the earliest defenders of the regulatory ideal. John Rawls' early work is a good example of contemporary RIT, as is Robert Nozick's libertarianism. Different as all their views are, each is concerned with specifying the ideal form of a just state and by so doing providing a framework for justifying policy in concrete situations.

Adam Smith, Karl Marx[15] and John Stuart Mill are some good examples of the CA tradition. Important contemporary comparative theorists include Amartya Sen and Kenneth Arrow. Each is concerned with specifying an impartial perspective and a

---

[15] That is, I think given two caveats; when his methodology was economic/historical in nature and when he argued that the proletariat perspective deserves to be seen in some sense as impartial, i.e., what is good for the proletariat would be good for all people.

28

methodology, as well as specifying what sort of concrete social and economic situations are preferable from their perspectives[16].

To see what philosophical benefit theoretical pluralism provides, it will help to focus on Amartya Sen's attack on RITs. In *The Idea of Justice*, Sen focuses his attack on social contract-style reasoning, which is only one type of constructivist reasoning in ethics. Social contract reasoning specifies criteria for the acceptability of an agreement or contract under ideal circumstances (and what counts as ideal circumstances should be constructively abstracted from a concrete level—more on this later). This ideal contract is then used to generate principles to organize the moral or political world under consideration. However, I will show how his argument can be refined to target the assumptions shared by constructivist meta-ethics in general. Before I do this, let me introduce the main target of his criticism: John Rawls and the contemporary social contract reasoning developed around *A Theory of Justice* and *Political Liberalism*.

1.2.1 Rawls and Social Contract Foundations

Rawls' original motivation in *A Theory of Justice* [1999] was to establish for political theory a basis that can do for it what set theory did for mathematics. The idea underlying his approach is that if we were to decide upon using principles *X* to organize society and we can show that we couldn't do better than *X* then a society based on *X* is justified. To this end, he develops the metaphor of a social contract in terms of what he

---

16 Not only is the CA tradition often tolerant of a plurality of correct but disagreeing principles, but in some case this conflict is theoretically front and center: Marx's approach for instance, may be characterized as explicating the nature of the conflicting principles of social organization in capitalist societies.

calls the 'original position.' The idea here is that we could use a decision to agree to a contract, suitably construed, as a basis for justifying the terms of forming and maintaining a society. If we could show how people in such an original position would agree to a best-case scenario of such social contract, the terms and conditions of their agreement could justify the resulting society.

However, the nature of how we set up the decision-procedure for our social contract must still capture what would count as a moral perspective, that is, if it is to have any claim to be a contract upon which we could not improve. Thus Rawls invents a 'veil of ignorance' that strips away any information from our contract negotiators that could give any undue advantage in the bargaining process. For example, knowing whether one is male or not would incline one to favor contracts that likewise suggest favorable outcomes for males or not as the case may be. Any sort of favoritism for outcomes based on social privileges or any other sort of lottery of birth would be cut out with such a condition. The notion of a moral perspective operative here is as a view taking persons as equal in fundamental ways deserving of equal consideration; here I focus on just one aspect of this equality of consideration, in terms of the fact that any knowledge of how anyone in particular would fare in a proposed contract is left in the dark.

As may be well known, Rawls came to the conclusion that we could not do better under these conditions than to agree to a society that maximizes liberties for each person compatible with the same liberties for any other and that an inequality of any socially, economically or politically significant good may only be allowed if it is to the benefit of those who end up among the worst off in that society. His reasoning for this conclusion

can be summed up in terms of a rational necessity for negotiators in the original position to take themselves to be in a worst-off position in whatever society they decide upon.

While his reasons for this averse attitude to the risks of ending up as a 'worst-off' are detailed, independently interesting and wide-ranging, I prefer to articulate them in a more general sense than he explicitly does but in a fashion that I propose is more in the spirit of both *A Theory of Justice* and *Political Liberalism*. Justifying society in terms of a sort of cooperation only makes sense first and foremost in terms of a risk-pooling arrangement in which we try to minimize the variability of economic activity and our exposure to violence. If we don't take this basis to be the purpose of distinguishing unjust societies from just societies, Rawls would suggest we are talking about different topics altogether. In other words, there are two things we must take to be primitive evils: catastrophic scarcity and violence. Since it follows that, rationally, we must seek to avoid these evils without thereby sacrificing anything of greater value, our negotiators in the original positions must assume the perspective of the worst-off, i.e., those most likely to suffer from the variability of economic activity or violence, and bargain accordingly. Of course once we are negotiating, other sorts of goods could become subject to similar reasoning: it is rational to pool the risk associated with catastrophic lack or loss of other sorts of goods if the benefits of doing so outweigh the costs.

1.2.2 Sen: Agreement isn't Acceptability

Sen, in his *The Idea of Justice* [2009], argues that contract-style RITs (what he calls 'transcendentalist institutionalism') all suffer from at least one of two problems:

they are either *infeasible* or *redundant.*[17] When he speaks of feasibility, Sen has in mind a

requirement for a certain sort of completeness. The incompleteness he accuses RITs of

should not be confused with the logical notion of theoretical incompleteness, i.e., that for

some sentence *p* the theory doesn't decide whether *p* or *not p* is true. Rather, the salient

sort of incompleteness refers to a theory failing to account for all the phenomena it should,

i.e., a "feasible" theory is one that accounts for all the relevant phenomena. Sen argues

that it implausible that any agreement (set of principles) specifying a perfectly just state,

or its institutions, could account for all the many, diverse impartial reasons that actually

exist for thinking of justice in one way or another.[18] This is what he means by the claim

that any unique ideal will be infeasible. More precisely, he claims that it's infeasible that

any one particular social contract could achieve unanimous assent from every reasonable

would-be citizen.

When Sen charges RITs with being redundant, his line of thinking is what one

would expect, given his terminology: if we need CA anyway to justify concrete policy

decisions, and if we can also use CA to justify the basic constitution of a state, then the

RIT framework is not doing any work. There are clearly two subsidiary claims involved

here. First, that RITs are of no help for solving particular policy problems, so a CA such

as the social choice theory Sen advocates is unavoidable. Secondly, that there is no

---

[17] According to Sen "with transcendentalism, I see two problems…the issue of the *feasibility* of finding an agreed transcendental solution…an exercise of practical reason that involves an actual choice demands a framework for comparison of justice for choosing among feasible alternatives and not an identification of a possibly unavailable perfect situation…this is the issue of the *redundancy* of [RIT]"[2009, pg.9]

[18] Moreover, the social contract tradition developed under Rawls posited an ideal contract, which everyone would accept and if it is implausible that such a contract could be constructed then such a RIT perspective is in trouble; or so goes the infeasibility complaint.

principled difference between justifying the choice of concrete policy and the choice of

what general policy should guide particular policy options (that is, the choice of the basic

constitution of a state).

While Sen's argument is careful and detailed, we need not cover his particular

attack on Rawlsian and other social contract theories. We can already abstract the more

general lesson from Sen's argument. To do so, keep in mind the constructivist

assumption: we assume that the notion of *acceptability* under ideal conditions is suitable

to play the role of 'truth' in ethics. Sen's infeasibility complaint points out that the set of

acceptable principles of justice is far larger than any one contract we could agree to.[19]

That is, the notion of *agreement* (to a contract) does not exhaust the notion of

acceptability (of a set of principles). As I want to refurbish his argument, there is an

analogous complaint against any particular specification of what counts as *acceptable*. In

other words, the class of principles we have reasons to accept is larger than any set picked

out by any RIT.[20] To be fair, I don't think I am claiming anything that is not already

implicit in Sen's criticism of a monist foundation.

There are two ways that I will take to ground the claim that there is no unique

self-consistent way to formulate what it takes for a moral or political principle to be

---

[19] Many of Sen's chapters are devoted to this claim; however the simplest and most helpful for our current purposes can be found on pg 57: "Indeed, plurality of unbiased principles can, I would argue, reflect the fact that impartiality can take many different forms and have quite distinct manifestations… [people's different] reasonings about the nature of a just society [can] reflect different basic ideas that can each be defended impartially (rather than being parasitic on vested interests)."
[20] Sen's infeasibility complaint is not unlike G.E. Moore's open question argument. More precisely, there is a more sophisticated notion indicated by the open-question argument that seems analogous to the infeasibility complaint. The idea is that any specification or discrimination of moral properties or principles cannot exhaust every possible reason we could take as action-guiding, morally speaking.

acceptable. First, the notion of *acceptability under ideal circumstances* can be made more precise and detailed in many distinct but equally justifiable ways. For instance, what should count as ideal circumstances can vary: in terms of what knowledge we would have; the time span of the contract; the operative; and whether some basic moral assumptions or sentiments (e.g. violence is bad) are presupposed such that we must allow some claims to have default justification.

The second way to try to ground the distinction between the generic and specific notions of *acceptability* is to mimic, *mutatis mutandi*, Sen's infeasibility complaint or some version of G.E. Moore's open-question argument. For instance, suppose we have some specification of the sort of reasons that should be taken to render some policy acceptable. We would have some formula like: reasons to choose *A* have *X* if and only if the choice to *A* is acceptable. Despite the general rejection of Moore's intuitionistic Platonism,[21] many people may still respect the intuition behind the open-question argument: that we can never be certain that there are no relevant properties outside of *X* that we should take as good reasons to guide our assessments of choosing *A*. In other words: But is it [really] acceptable for reasons to choose *A* to have *X*? For example, let *X* be some formulation of *impartiality*. On this supposition, the reasons to choose *A* are impartial if and only if the choice to *A* is acceptable. But, it still seems intelligible to ask: is it morally acceptable for reasons to choose *A* to have that particular formulation of *impartiality*. Is it not possible that we could do better, morally speaking, with some other formulation?

---

[21] The most straightforward reason for rejecting Moore's proposal I take to be is Bernard Williams' [2012] demonstration that Moore equivocates on the attributive and predicative properties of an adjective (pg 38-42.)

A related problem is that any correct characterization of what counts as a good reason must on some level be attentive to contextual factors. If this is the case, the implausibility of enumerating every possible correct characterization is due to how contextual factors are generally understood as having an open-texture. As is well known, it is impossible to enumerate every characterization of a phenomenon if under conditions of open-texture.[22]

The redundancy problem needs no substantive changes for it to apply directly to any specification of what counts in general, as a constructively acceptable characterization of an ideally just state.

1.3 The Solutions: Moral and Political Pluralism

The Ought/Must Distinction and Moral Pluralism

At first glance it is easy to see why those who feel inclined to agree with Williams' assessment of the plurality of best moral practices might also buy into anti-foundationalism (or more precisely, a non-systematic foundations) of some sort. In fact, particularists (and possibly moral sentiments theorists) of the more careful and subtle variety could try to use a version of the *obligations as best options* idea to substantiate the claim that understanding which moral principles are relevant to a choice is unrelated to making an appropriate judgment regarding what moral properties obtain in that situation.

---

[22] The structure that I argue captures a related notion of this sort of open-texture of moral reasoning is the *choice* parameter. In chapter three, I argue that the moral properties of an action obtain only relative to a contrast class of unselected options. Since we can always specify different or additional members, i.e. options, of the contrast class, the moral properties of an action demonstrate an open-texture. A similar structure should be able to be used to capture the open-texture of good reasons.

But this inference is sound only if foundations need to be systematic and if a systematic approach needs to be self-consistent given the right logic. While a non-systematic foundation might be a plausible and interesting project, it is not the current one.

Nonetheless, we have a clear criterion for what we are looking for in a suitable systematic foundation: a clear and principled distinction between ought and must. If we are to have a systematic foundation that takes the distinction between ought and must seriously, we need a system of many equally correct moral theories. With this notion, it is possible to have cases with many best options as distinct from when there is only one best option all theories posit. And from the discussion regarding Sen's attack on RIT approaches to justice we find something similar: we can't use only one RIT to account for all we should when deciding what is unjust. The natural idea I explore from here on out is to allow for a plurality of RITs in what I call a PRIT meta-ethics.[23]

1.3.1 *Infeasibility, Redundancy and Political Pluralism*

The infeasibility and redundancy problems are powerful reasons to reconsider the regulatory ideal approach when such approaches underlying reasoning rely on the notion of *agreement*. Moreover, as I have revised Sen's argument, there are good reasons to reconsider the RIT approach at large, given that the underlying reasoning of such approaches may likewise rely on any one particular notion of *acceptability*. However, while I grant the force of these arguments, I think we need to be careful to draw the right conclusion. Instead of rejecting the regulatory approach *tout court* in favor of the comparative, as Sen urges, I suspect that the infeasibility complaint uncovers an

---

[23] I pronounce this acronym: pree-tee. ... as in a pretty nice meta-ethics.

unwarranted assumption made by defenders of the regulatory ideal framework. As I shall argue, this assumption is unnecessary and a view more plausible than either the standard RIT or the CA results if we delete it.[24] The assumption to be rejected is that liberalism must be consistent to be correct. In other words, our fundamental principles of justice may disagree. And by happy coincidence, this idea fits nicely with what we uncovered from the discussion of William's *ought/must* distinction.

I am not unsympathetic to social choice theory, nor would I reject Sen's demand that there be a suitable mechanism for deciding when concrete choices must be made. Rather, I will argue that his infeasibility argument fails to take into account a promising option. The key idea is that the infeasibility problem misses the target if pluralism about ideally just institutions, procedures or policy is taken seriously.[25] In fact, RIT pluralism (i.e., PRIT) takes the diversity of impartial reasons that are the key to Sen's objection as its starting point for thinking about justice. By the 'diversity of impartial reasons' I mean that no one theory of *acceptability* can account for every member in the class of principles we have good reasons to accept.

Proponents of an RIT generally assume we need one uniquely correct account of perfectly just institutions to serve as the regulatory ideal. Sen shares this assumption, at least implicitly (i.e., the infeasibility argument has force only if there is at most one regulatory ideal.) But one might suspect that PRIT has traded frying pan for fire. For if

---

[24] However, to argue this point regarding the CA project is beyond the scope of the current project. At the very least, the current project should set the stage for such a comparison down the road.

[25] This pluralistic approach looks like it could generalize to other domains where some sort of open-question argument seems legitimate. For example: norms of assertion, logics (i.e., are there logical truths not included in this logic?), norms of theory selection, and scientific methodologies in general. However, while interesting, pursuing the question of whether this is so is outside the scope of this project.

there are two regulatory ideals relevant to a policy choice, but each ideal selects a different option, then we have an action-guiding problem. Perhaps worse, each option is recommended as ideal in the circumstances. There is a far more problematic possibility waiting the pluralist. Since in the larger picture we are assuming that anything prescribed by an ideal is prescribed, there is at least one logical rule that entails disaster. The logical rule of "explosion" (i.e., that everything follows from a contradiction[26]) would seem to entail that anything and nothing is just if we accept inconsistent principles. It seems that we have traded an arguably difficult problem of incompleteness for a much worse problem of over-completeness or the failure of action-guiding.

1.3.2 *Deontic Logics: Moral Concepts as Logical Words*

In order to discuss how I think we should deal with these issues, it seems appropriate to explain some of the reasons for the formal methodology of this project. However, let first me outline where I deal with the specific issues raised already. To address the concern that pluralism implies an inability to be action-guiding, I examine in Chapter 2 a few ways philosophers have looked at the phenomenon of *conflicting principles*. In Chapter 3, I develop the underlying structure for a relevant deontic logic that allows us to represent *trumping, conflicting principles* and *justifiable choices* in an illuminating way. It is my contention that by structuring these concepts in this way we will be able to better account for our actual practice. In Chapter 4 I shall present some technical tools that show that a set of inconsistent regulatory ideals does not imply that everything is permitted. That is, I will develop a paraconsistent deontic logical system

---

[26] This rule is valid, for example, in both Classical and Intuitionistic logic.

(i.e., one that does not validate explosion) that also avoids pitfalls that undermine some other deontic systems. These pitfalls include committing the naturalistic fallacy and ignoring trumping relations.

I will also try to show that action-guiding under-determination isn't the problem it is commonly thought to be (and that Sen assumes it to be). I shall argue in Chapter 3 that any regulatory ideal must specify the terms of fair cooperation, and this will include a specification of the equilibrium selection mechanisms appropriate for coordination, i.e., how we decide on who does what to cooperate must be identified. If I am correct, the appropriateness of these mechanisms should not be independent of the theory in question, i.e., *how* we decide who does what must be fair as well. It follows that the fairness of coordination is theory-dependent whether there is only a single ideally just state or if there is, instead, a plurality of ideally just states. However, this is a useful fact for the pluralist. It opens up the possibility that when two distinct policy options are selected by different regulatory ideals, we can in some cases resolve the action-guiding problem. In some cases, we will be able to decide on which policy to implement by using an equilibrium selection mechanism that both theories accept. I will illustrate how this works when I have the right representation of trumps reasoning, thus showing how PRIT could answer Sen's redundancy objection.

Now, let me turn to explaining the methodology and defending how I will deal with these issues. The plausibility of a pluralist foundation for liberalism depends on whether we can find (i.e., construct) an intelligible notion of prescription or its cognates for such an approach. In the most basic analysis, pluralism, arguably, solves major problems with political liberalism; the problem that arises is that with a plurality of

liberalism, it's difficult to make sense of how several theories can together guide the choice of policy. The method I propose we use to address this issue is a type of deontic logic.

Deontic logic may be characterized by the idea that we can treat the concepts of moral *obligation* and *permission* as logical words [Scotch & Jennings 1981; Hare 1964]. For example, let's say we have a description— "a red apple is on the table"; we can transform this statement into a prescription by prefixing it with *"there should be"* resulting in "there *should* be a red apple on the table." Thus, insofar as the concept of obligation can be articulated in terms of a *moral* 'should' or 'ought to', it is apt to be described as a sort of modality and perhaps formulated as such in logical terms.

Like other logical words, such as *"and", "or"* and *"not"*, the idea is that we should represent obligation and permission as operations on well-formed formulas. The starting point in deontic logic is that we represent the concept of *obligation* as a sort of necessity operator. Obligation, so understood, is what is *morally* necessary. However, in what follows I do not use the standard method in modal logics, where we use some sort of universal quantification over suitably related *worlds* to represent necessity. (In the deontic case, these worlds could represent what would be true[27] in morally satisfying or, perhaps, ideal situations.) Instead, I defend an analysis of *obligation* independent of possible worlds talk and thus turn to what is called *neighborhood semantics* for the truth conditions for obligation; but more on this in Chapter Four. In this chapter, I begin the overall task of showing how structuring *obligation* in terms of moral necessity is quite a

---

[27] Talk of truth at morally accessible worlds is a heuristic for now, which will be developed into the more refined notion of *designated value* in chapter 3.

complicated formal exercise if it is to do justice to features found in actual lived moral experience.

1.3.3 *Choices*

There are some considerations that call for a refinement of this basic intuition, i.e., that prescriptions are suitably modified descriptions. The basic idea is that not just any sort of description is appropriate for such transformation, but only those that express a *choice*. To make the notion of a choice amenable to formal treatment, it will be useful to be able to treat choices as mathematical objects of some sort. I will therefore treat a *choice,* on this account, as a (always finite in the examples we deal with) countable list of *options*. We will informally understand the options to be such that an agent confronted with a choice must select an option, and to be constrained by the rule that once an option is selected no other listed option can be, at least in that case. As a heuristic, morally obligatory choices can be thought of as choices where no other listed option is morally better than the selected option.[28]

To illustrate the sorts of descriptions I have in mind here, instead of merely 'a red apple on the table' we would need to also describe the other relevant options, such as 'a green apple on the table' and perhaps also 'a red apple on the floor'. A *choice* then would be to make it true that a red apple is on the table instead of a green one, or rather to place the red apple on the floor.

---

[28] This heuristic seems to indicate that a (moral) preference ordering could generate such a ranking. Though I will not pursue this, one promising way to construct such moral perspectives is to posit ideally moral agents, saints or vampire slayers, and order choices in terms of such saints' preferences.

Why *choices*? I think there are many good reasons for this proposal, some technical, others more informal. For now, let me introduce some of the less technical concerns. Suppose it is morally objectionable to choose to lie to a reporter instead of telling the truth when the truth will likely result in you losing a two million dollar investment. However, if we pay close attention to the fine-grained structure of choices, we can represent a change in the moral characteristics of lying. To do so, we change the available options. Here's how: it seems morally obligatory to choose to lie to a potential heart attack victim instead of telling the truth or remaining silent. Given a basic training in CPR etc., one should know how telling the truth, i.e., that 'yes, I think you are having a heart attack' or even just staying silent on this fact, will likely result in a worsening of symptoms or death. While it is true that, linguistically, we can transform any description into a prescription, my claim is that such linguistic phenomena is, in this case, conceptually misleading. In Chapter Three I go into more detail on the nature of what I call *Contrastive Principles*: that principles, obligations etc. bind choices, not the constitutive actions nor propositions, classically understood.

One further example before we move one.[29] There is a flawed way of understanding human rights that, if we were to use contrastive principles or *choices* instead, we could say what we actually mean far more simply and accurately. I think the idea that a person could do something that thereby revokes or forfeits some (or all) of her human rights is incoherent. Of course, there is a seeming consensus disputing my claim, i.e., regarding how people could in fact lose or forfeit human rights [Lippke 2002; Ward

---

[29] An earlier version of this argument shows up in Holukoff [2006].

and Birgden 2007],[30] and the brief argument that follows could hardly be the last word. On the other hand, I propose that we can say what we really mean with *choices* and to do so without any of the conceptual gymnastics that seem to be needed to make sense of how it is possible for someone to lose her human rights.

The usual way to understand human rights is with caveats—one can lose or forfeit her human rights by doing something bad. This is natural only if the reasons for having human rights also pick out a category of agents and membership in this category isn't permanent, i.e., it can end before biological death and perhaps even be resumed. There are three problems I want to highlight with this idea, i.e., that membership isn't permanent (in the above sense) in an 'eligible for human rights' category. All three problems are due to the fact that any right, including human rights, should be an instance of a general rights formula. By the *General Rights Formula*, I mean that: x has a claim R against y in virtue of Z, where x is the right holder for the right R and y is the respondent (the agents or institution that must see to it that R for x) and Z is the basis or reasons for why x has R against y. And of course, it's implausible that human rights don't take this general form.

It turns out that if human rights are an instance of this more general rights formula, it seems that membership in an 'eligible for human rights' category can't fail to be permanent in the way required by the caveat approach. The first two reasons I will not argue for (but rather give the benefit of the doubt to the caveat approach): how we specify Z for caveat-style human rights is likely either too broad, i.e., it implies we includes things in an 'eligible for human rights' category we shouldn't, or too narrow, i.e., it

---

[30] To be clear, the authors cited dispute this consensus as do I, but in different respects.

excludes persons it shouldn't. But as I suggested, I want to suppose that somehow we do get Z right; if so, there is a more interesting problem that awaits us. If Z is neither too narrow nor too broad, but is compatible with the caveat approach, it implies that people who lose human rights aren't human or that the various relevant biological markers that should be coincident with reasons for human rights fail to be so.

I will assume without further argument that claiming that 'people who lose human rights stop being human' is unacceptable. On the other hand, it would seem that the various stories for how the relevant biological markers aren't coincident with reasons for human rights are all a mess—for an easy target, reciprocity is a terrible basis for a comprehensive moral category as it ignores the most vulnerable and hence those most in need of the protection afforded by human rights.[31] Of course, I promised to give the benefit of the doubt regarding how we specify Z, i.e., in a way that is neither too broad nor too narrow, so I will put these concerns off to the side so we can focus on the real issue. There are various biological markers that should be coincident with reasons for human rights and any story that claims to justify them failing to be so is inappropriate. With the stage set, let me explain in more detail what I mean by 'relevant biological markers.'

Consider Mary. Mary is a human, and, however one feels compelled to articulate them, she has all the features relevant to her claiming *human rights*. In fact, any correct formulation for the grounds of human rights, i.e. Z, must determine a set that is implied by being human. This is not to say that 'human being' is a moral category; that is, we can

---

[31] And the expected response fares no better. If it's the *potential* for reciprocity that matters, we run into problems in the other direction: 'potential x' is a very broad category indeed.

say, of some human right—say *R*—that all humans have *R* because every human has *Z* (or belongs to *Z* etc.). To paraphrase Brian Orend [2002], human rights are founded upon reasons to treat every human in the relevant fashion (pgs. 19, 34.) And of course, Mary *being human* is not such a reason; rather it is that *Z* picks out Mary, just like any other human, which matters, morally speaking. In other words, it is a coincidence, but by design, that being human biologically (together almost certainly with other biological features, such as having a brain, being sentient perhaps, and so on) implies belonging to a *moral* category that thereby confers upon its members certain rights. But it is tempting to think that *Z* has loopholes, by which a person may satisfy some condition whereby the reasons to treat all humans in the relevant sense no longer hold, at least for this person. Here's why.

Suppose we have the relevant sort of very good reasons to not place Mary into a dangerous situation against her will. And such good reasons turn out to apply equally to any human being: that is, we should not place any human in situations of unreasonable risk, i.e. we have no right to do so. But alas, Mary shoots Jack with a shotgun for telling her where to put her apple. In societies such as ours, prisons tend to be very dangerous places. If Mary has a right to not be placed in a dangerous situation against her will, it appears that she has, like all humans, a right to not be sent to prison against her will.

This seems all too fast: prisons are a fundamental affront to human rights? Now, as suggested, there are all sorts of stories trying to show how the relevant human rights stop applying when certain crimes are committed, but all such stories are unnecessary and, if I am right, incoherent. Such stories seem to implicate that when we refrain from applying human rights to any particular human this is equivalent to denying that she or he

45

is human or that the relevant biological markers that should be coincident with the reasons for human rights fail to be so. To see this, we can formalize the set membership talk from above by considering a set-theoretic deduction.

We use *H* to represent the set of all human being, while *Z* and *R* are as above. We can read *xRy* as '*x* has a claim *R* against *y*'. And as promised, taking this approach will imply a *General Rights Formula: x* has a claim *R* against *y* in virtue of *Z*. If this general rights formula is applicable we have the following inference that shows that denying a human right to a human is incoherent.

$$\forall w, x, y \, (w \in H \rightarrow w \in Z) \, \& \, (x \in Z) \rightarrow xR^n y \vdash \neg xR^1 y \rightarrow x \notin H$$

This inference obviously requires a more detailed interpretation and defense than can be given here. But the idea is that anytime a human right is denied, the humanity of that person is denied *since membership in Z must fail* and by *modus tollens* so to does membership in *H*, i.e., being human. Of course, there is the possibility that what is going on here is that we are slipping between a generic claim about humans, generally, having said rights and confusing this with a universal claim, i.e., that all humans have said rights. But taking this route suggests, implausibly, that the general rights formula isn't applicable to human rights. But even if this is so, it then becomes strained to see just what work using biological markers, i.e., being human, as a shorthand for assigning moral membership into this class of rights is doing.

Practically speaking, this suggests that caveats to *Z*, such as 'as long as you don't commit certain offences' would mean that *H* doesn't imply *Z* after all, and the fear is, there be dragons. The practical value of human rights on this account is that appeal to

them shuts down further discussion, so to speak. No government official or tyrant can spend enough time to dream up some scheme by which someone no longer deserves to be treated with basic human dignity, for example, and this because we can simply point to the relevant biological markers.

Again, I don't pretend that this brief line of argument is conclusive. The hope is that readers might now share some of my suspicion towards the caveat approach and perhaps find a simpler and, I propose, a far more accurate way of understanding human rights more attractive. Simply put, it's just a fact that people don't have a general right, human or otherwise, to not go to prison when they are rightly convicted of the relevant crimes. It is a virtue of my system that, by using contrastive principles and *choices*, we can see how we missed such an obvious response and how to fix the problem.

In my system, it is the content of the right itself that is at fault for such a bizarre result (i.e., that there is a human right against incarceration). The clue perhaps is the word 'unreasonable' in 'we should not place any human in situations of unreasonable risk.' The appropriate response seems to be that the dangers of prison are not unreasonable, that is, *given* that Mary's shooting Jack constitutes murder. It is by glossing over the *conditionality* of the content of rights that give us such bizarre results as seen above. On my account, we represent rights as an obligation for someone to make a particular choice. The choice in this case: Should we *leave someone alone instead of incarcerating her for murder*? When the issue is clarified by thinking in terms of choices, it turns out there is no human right for governments to leave you alone when you murder someone. That is, any legitimate take on *Z* won't justify making the choice to leave someone alone instead of incarcerating her when rightly convicted of murder, all else being equal. Analytically,

choices explain what's going on when we correctly (I assume) limit rights in such cases; there may be other analysans, but the simplicity and technical power of *choices* turns out to be remarkably well-suited for our current purposes.

1.3.4 *On Using Formal Methods in Moral Foundations*

However, the further question remains: why use formal methods to investigate the notion of moral necessity? All I have done so far is to show that we could do so and illustrate some of the utility in so doing. On the other hand, I am not unaware of Aristotle's injunction against precision in moral theory.[32] He argues that we should never aim for more precision than we should expect from different fields of inquiry. When articulating doctrines in different fields, precision can be counter-productive when accuracy is better served by remaining more vague. He found this to be especially the case with moral inquiry where he held a large place for reasonable differences in moral judgment. In fact he contrasted the work of a mathematician with that of an ethicist as an example of when we should expect or reject a high level of precision.

Nonetheless, the trade-off between accuracy and precision in ethics can be moderated by the development of better tools of precision. I propose that, even if Aristotle's injunction was at one time wise advice, its appropriateness needs to be reassessed; given the [massive] progress we have made in designing artificial languages and their formal techniques, it's not unreasonable that some updating is called for. His point was that precision is limited by the nature of the subject matter and I don't deny that this is true, but only to some degree. I propose that better tools can extract more detail

---

[32] Aristotle, *The Nichomachean Ethics.* Book 1 Chapter 3, first paragraph.

and illuminate finer distinctions than more coarse-grained techniques, for nearly any subject matter. And on the flip side, better tools may allow us to safely ignore some details and distinctions when called for (more on this later). I think we can do better than what Aristotle proposed, i.e., leaving the nature of conflicting principles up to legitimate differences in mature, trained, and considered judgment or sensibilities. I hope my results will help make this case but I also think there are good reasons to expect this outcome.

On this note, it is worth a short tangent to see how in particular using logical machinery can advance our understanding, even in a domain like moral theory. When we set up a formal calculus and it's associated formal semantics, in order for it to be "the logic of " some domain, we hope to build the machinery in such a way that it captures the concepts we are interested in investigating. And as I'm sure even Aristotle would agree, this is more art than science. Once done, though, it is a familiar fact that formal tools can enable us to reveal things that are simply not obvious, and in some cases is seems safe to say, would never have been discovered, had the formalization never happened. For example, the informal understanding of notions such as provability or truth has changed dramatically thanks to the investigation of the formalized versions of these important notions. The hope is that we can do the same for "obligatory," "permissible," and the rest.

Of course, there remain philosophical questions about the relationship between the concepts of truth or provability and their formalized versions. For any particular formal result, it is always open to ask whether the result teaches us something about the concept under investigation, or if it only teaches us that the formal analogue is different from the concept we care about—consider, for instance, Tarski's [1956] well-known result about "no language containing its own truth predicate." However, this process is

fruitfully viewed as one of pursuing a sort of reflective equilibrium, and both adjustments

to the formal machinery to better reflect the actual concept and revision to our thinking

about the nature of the original concept are both on the table.[33]

With this said, what I will be trying to do in this project is provide a formalized

version of key moral concepts, with the hope of learning some useful lessons about those

concepts, ones that would not be easily learned if the formalization were not undertaken.

But, as should be clear from the discussion above, attention must first be paid to the

question of what to build into the formal system I will investigate. As I have already

suggested, I think standard deontic logics won't do the job. To get a sense of what we

need to do differently, it is necessary to do some preliminary philosophical investigation.

And to get a sense of the subtle and nuanced deviations a pluralist and constructivist

notion of *obligation* undergoes, let me briefly survey some of these philosophical issues.

While it has many virtues, the meta-ethics I will look at must reject many

common theoretical assumptions regarding the concepts of *principle* or *obligation* and

their cognates. As mentioned in the Introduction, I show that neutrality towards the

existence of moral dilemmas (i.e., taking the possibility of moral disagreement seriously)

renders the inference "*ought* implies *can*" invalid.[34] Moreover, in the formal language I

---

[33] Moreover, even if a formal version diverges significantly from its informal analogue, such a contrast could prove illuminating in surprising ways. For example, such divergence may call for explanation, which in turn may point to some novel significance of a feature of the informal concept. I am thinking here of the divergence between the material conditional and the ordinary language locution of "if...then..." and say, Grice's [1975] notion of conversational implicature to account for it.

[34] I also give a non-technical argument against "ought" implies "can" in chapter three. It is based on intuitions that we should not be able to jury-rig the order of actions so as to immunize ourselves from living up to our obligations. It is closely connected to technical issues however in CTD structures.

develop for PRIT, *RD-choice*, I prove that it is only in ideal situations that *ought to choose X* implies *it is permissible to choose X.*

Specifically, in 4.68 and 4.69 I argue that any conception of *obligation* that implies *can* or *permissibility* is inapplicable, all things being equal, to any world like ours. With the system I have developed, it will be possible to distinguish between principles and obligations that can guide action at an ideal level of abstraction from the mess of moral properties we tend to see in concrete situations that end up insufficiently guiding action for people like us. I also prove that, given neutrality towards the existence of moral dilemmas, obligation and permission are inter-definable only under ideal conditions.[35] To be more precise, I show that it is *must* rather than ought that implies *may* (technically, *must* implies what deontic logicians call *"weak permission"* but more on this later).

At first glance, this may seem a radical modification of *obligation* and so may be thought implausible; I shall argue that closer attention paid to our best practices reveals that my proposed theory accords better with such concrete touchstones. The question then becomes: Is it reasonable to cling to a conception of obligation, when a more nuanced conception fits our practices better, even given how obviously correct it feels? In what follows I hope to demonstrate that a correct account requires a far more subtle and careful relationship between obligation and action-guidance than is commonly supposed. This holds even more so in political discourse.

---

[35] Following standard practice in deontic logic, technically such notions of *permission* are called 'weak permission' in contrast to the notion of *strong permission, which* is an exception to an obligation under specific conditions. CTD structures will be used to regiment the notion of *strong permission*, i.e. as a sort of sub-ideal best practice.

As promised I can now explain how we can have a non-theory (since theories can be thought of as consistent sets of propositions) but systematic approach to meta-ethics. If we can use a deontic logic, i.e., a formal system, to represent the claim "*obligation* is related to *must* as *best* is related to *only*" then I have proved that it is at least possible to have a systematic moral foundation that can account for Williams' starting points in concrete moral experience. Thus while logics are theories and thus consistent themselves, we can apply logics to domains that aren't consistent. But, since we can represent what is going on in terms of a single logical system, we have what I hope deserves to called a systematic foundation. It just turns out that all the principles etc. that we represent with the logic are themselves not consistent, and thus, not a theory. In fact, one of the main results of the formal presentation is that we can justify the claim that liberalism itself is an unrealizable plan, not that it just can allow for reasoning regarding unrealizable plans—see in particular 4.64.

1.4 *Constructivism, Pluralism and Prescription*

As I have laid out the project, the style of meta-ethics I'm looking at deals in the construction of an ideally just society (in our case, ideally liberal). The issue that pluralism raises, in large part, is how must our understanding of the doctrine "whatever is prescribed by an ideal, all things being equal, is obligatory, a principle of justice etc." change?

To see why it is reasonable to suppose that changes must be made, let us assume that this characteristic inference is correct: *what a correct moral theory or ideal prescribes is prescribed*. This seems pretty reasonable, or it's hard to see what the

correctness of the theory consists in. However, if pluralism means that we will be confronted with more than one correct moral theory, the prospect of being obliged to more than one, incompatible thing looms. Conflicting obligations raise a whole host of problems, so at least two questions arise at this point. First, is the cure (pluralism) worse than the disease? And secondly: why isn't pluralism just a code word for "we haven't got liberalism right yet"? I will answer the second question next. The project at large is intended to answer the first.

1.4.1 *Constructive Foundations and the Burden of Pluralism*

All this talk about methods for understanding moral concepts in a pluralist-friendly manner may seem to be putting the cart before the horse. Here's what I mean: just because some important philosophers are talking about or have taken a pluralist turn in the face of persistent problems doesn't mean that pluralism is the right way to go in the foundations of liberalism. After all, we might just have gotten things wrong or be working with incomplete theories so far. Instead, perhaps we ought to hope that in the foreseeable future we will discover one uniquely correct theory that captures everything we thought mattered in previous attempts. In this section I will argue that this hope is misguided, at least if we are proposing constructivist foundations for liberalism.

Once we pay close attention to what makes a theoretical principle suitably constructivist in nature, we will see that justifying a constructivist approach in any foundational domain generates a *prima facie* case for pluralism in that domain. I contend that it is implicit in the nature of constructivist foundations (for any domain) that there are good reasons to be suspicious of any claim to some uniquely correct approach. If this

argument is cogent, the burden of proof is on any constructivist to show why she is not also a pluralist.

I will also briefly look at a plausible premise that, if added to the argument, suggests that in order to justify constructivist foundations, one must also justify a pluralist foundation. However I am less sure about this second, extended version of the argument; while the additional premise seems plausible, giving a fuller defense of it is part of a much larger project. A quick note: I don't mean that to justify constructive *methods*, one must justify pluralist foundations; some classical mathematicians are nevertheless fans of constructive methods, but merely because they give better details or more informative proofs (for example) rather than because they think they are fundamentally correct. Instead, in the argument that follows, I propose that any type of *foundations* that can deserve the title of *constructive* is also, all things being equal, very likely a pluralist foundation.

It will be useful to approach this argument by answering two questions. First: what makes a principle, method, or foundation *constructivist*? Secondly: assuming that a genuine plurality in foundations requires interesting, non-trivial disagreement between foundational theories, what is the connection between constructive foundations and interesting, non-trivial disagreement?[36] As a first pass, the answers I have in mind are as follows. For something to count as constructive in the relevant sense means being able, at least in principle, to trace the grounds for any true claim back to the basic sources of

---

[36]In putting the matter in terms of such disagreement, I follow [DeVidi 2012.] I discuss DeVidi's criteria for a philosophically important pluralism in more detail in Chapter Two. At this point, I will only argue that the pluralism I have in mind have these properties. Later, I will show how and why these properties matter in the relevant meta-ethical issues.

evidence underlying it—there are no epistemically inaccessible truths. But as I shall try to

show, this means that a correct constructive claim is relative to a starting point. As I will

suggest, this is naturally thought of as relativity to a *choice of* starting point. Not just any

starting point will do, i.e. our choice is not unconstrained. But there is little reason

(notwithstanding a commitment to some "mind independent reality," which alone makes

true statements true) to suppose that there's only one correct starting point. And of course,

believing that there is some such reality undermines the commitment to constructivism in

the first place. As we will see, more than one correct starting point will make for

interesting, non-trivial disagreement and thus a pluralism worth having.

This lesson is especially easy to see in the mathematical case, where

constructivism is better worked out than in any other domain in philosophy. To answer

these questions in more detail, I will develop insights drawn from the Italian school of

intuitionistic mathematics, most notably Giovanni Sambin's *minimalist foundations*.

Sambin's minimalist foundations takes *type theory* (which I will explain in more

detail below) to be conceptually prior to *set theory*. In brief, in Sambin's minimalist

foundations the definition of a *set* is not given by *set-theoretic* axioms but is worked up

from elements of type theory. In particular, in Sambin's theory *subsets* are not, in general,

sets (contrary to the standard ZFC axiomatization of set theory).[37] But not any sort of

type theory, but only a constructive version will do. Following Per Martin-Lof, Sambin

and Silvio Valentini [1998] explain what is at stake for an intuitionistic (i.e., constructive)

type theory (and thus for the possibility of a constructive foundation):

---

[37] A simple constructive rationale for this denial of set-hood for subsets: defining a
subset on a set is impredicative, and thus non-constructive.

The form of type theory is that of a logical calculus where inference rules to

derive judgments are at the same time [type] theoretic constructions, because

of the ["formulas-as-types"][38] interpretation. The spirit of type theory –

expressing our interpretation in a single sentence- is to adopt those notions

and rules which keep total control of the amount of information contained in

the different forms of judgment (pg 3).

In general, from an intuitionistic perspective the truth of a proposition (say P)

depends upon "the existence of a verification of" P (pg. 2). In type theory, talk of *truth*

and propositions is replaced with (in mathematical terminology) a *judgment* that $p \in P$,

that is, a judgment that some term, *p*, is of *type P*. An intuitionistic version requires that

such a judgment "explicitly exhibits a verification" that p is of type P. Paraphrasing Greg

Restall [2008], type theory allows us to 'name' proofs by keeping track of the terms used

throughout the "consecution" (or proof) (pg 127). Thus the judgments in type theory

record each step (i.e. rule and assumption) used to verify some conclusion. Not only does

a type-theoretic presentation of a proof tell you what was inferred from what; the

conclusion also explicitly contains all the assumptions, premises and rules used to derive

it.

On this account, the judgment that "*A* is a set" explicitly encodes the specific

procedure for delineating the elements of A (and, since the procedure is constructive, this

---

[38] The original primarily referenced the "propositions-as-sets" interpretation, but in keeping with [Sambin & Maietti 2010] proofs-as-programs interpretation, I used the other terminology.

procedure is often required to satisfy a further condition, such as recursiveness); Sambin

and Valenti write:

> In fact, the rules for primitive types and for type constructors are so devised
>
> that whenever a judgment *A set* is proved, it means that one has also complete
>
> information on the rules which describe how canonical elements of *A* are
>
> formed…Summing up, we see not only that [intuitionistic] type theory is
>
> inspired by the principle of control of information, but also that the same
>
> principle should be at the base of any coherent treatment of sets and
>
> propositions, if it has to be both intuitionistic and free of wastes (pg 2).

As it turns out, in their theory of subsets, a judgment that "*B* is a subset of *A*" doesn't

necessarily encode a specific recursive procedure for delineating the elements of *B*;

instead, the judgment encodes a specific propositional function on the relevant set (pg 7).

And without delineating the elements of the input domain, the function can't be used to

recursively populate the subset, i.e., it's impredicative.

However, there is such a thing as too much detail. In Sambin and Valenti's terms,

orthodox type theory is "a burden when dealing with the synthetic methods of

mathematics, which 'forget' or take for granted most of the details" (pg. 2). Consequently,

the goal is to develop "definitions and rules which 'forget' some information, and thus

allow a higher level of abstraction, which can make type theory more handy and suitable

to work out (intuitionistic) mathematics…" (pg. 3) However, for it to be constructivist,

properly speaking, we must be able to later recover what was "forgotten".

It can be helpful to focus on how we can take the notion of *forgetting details* as a

metaphor for *abstraction*. Similarly to how David DeVidi makes the point, abstracting

away properties of objects is a matter of forgetting those properties when formulating "principles that, as we might say, govern the phenomena in question." [DeVidi 2012] The stricter the principles, the finer grain of distinctions we can make. On the other hand, we must be careful that we are not merely changing the subject matter by changing how much or what we abstract away. An (overly?) simple way of looking at this is that age-old maxim: make sure you compare apples with apples. There is nothing amiss, all else being equal, if we compare apples with oranges when what matters is *being fruit*. Alternatively, if what matters is *being baking apples* the criteria we use to compare apples should reflect what goes into making an apple suitable for baking. Clearly, comparing fruit is a different subject matter from comparing baking apples; thus abstracting away too many properties surely results in a change of subject matter. But comparing sweetness and comparing ripeness may be seen as two different bases of abstraction in the subject matter of 'when apples are ready to eat'. In other words, different principles governing 'when apples are ready to eat' can be obtained by alternatively forgetting one property or the other (i.e., sweetness or ripeness). At this point, we have different ready-to-eat principles; but where is the interesting and non-trivial disagreement?

Given just these principles (i.e., the ripeness-principle and sweetness-principle), there may already be disagreement about which apples are ready to eat. But this possibility (i.e., one principle identifies an apple as *ready-to-eat* but the other doesn't) would seem to suggest something went wrong with the choice of one principle over the other. After all, a principle formulated with both properties in mind (i.e., those apples belonging in a union of *ripeness* and *sweetness*) would be preferable, all else being equal.

But it is sometimes hard to tell when *all things are equal*; this fix is sometimes more trouble than may first appear. In the case of intuitionistic mathematics, it is impossible to combine a principle that the collection of all subsets of a set is a set (the *Powerset Axiom*) with the principle that every total relation has a non-empty section that contains a function (the *Axiom of Choice*) without proving the law of excluded middle (and thereby become incompatible with intuitionism) [Bell 1988; Maieti & Valentini 1999; Maietti 1999]. Given different starting points, these principles both encode something that is desirable in, for instance, the notion of "well behaved concept of set", and so are analogous to our imaginary case of disagreement about which apples are ready to eat. However, trying to come up with a stricter notion of well-behaved concept of set that incorporates both requirements doesn't work, because a classical notion of set is one that is, for the constructivists in question, beyond the pale—rather as though apples that were both sweet and ripe were also, *ipso facto*, toxic. On the other hand, let's say that the single-property constructive principles agree with each other every time about which apples were ready to eat…what sort of disagreement can still arise? One sort of disagreement is fairly trivial. The disagreement I have in mind here is when the *verificational* content of the judgment, i.e., the terms in which the conclusion is *proved/confirmed*, differs. In this case, there is no disagreement about which apples are *ready-to-eat*, only in the tests recorded as used in the different judgments: in one judgment ripeness tests are explicitly recorded and in the other judgment we would have sweetness tests. This disagreement is trivial: *of course* different tests would be recorded corresponding to the different *properties* characteristic of the principles. Taking different

paths to the same destination, even if one is more scenic (or informative) than the other,

doesn't really amount to much disagreement about where you are now.[39]

The second sort of disagreement is far more interesting. This sort of disagreement

again assumes that the principles abstracted from the same 'reality', phenomena, or basic

notion nevertheless disagree. In other words, in a given domain or subject matter multiple

principles may equally apply but still disagree in some other manner. The question,

though, is how that could be: if they begin from the same "reality" and merely forget

different parts of it, how could the different remainders disagree? This possibility

suggests that the disagreement we are looking for occurs down the road, so to speak.

DeVidi writes that this possibility requires "that abstraction yield principles that, when

stated in properly abstract, i.e. formal terms, are correct when we reinterpret them in the

original domain. But once abstracted and formalized, mathematical practice involves

extending principles as far as they will go, or at least as far as we can push them. This is a

process that fairly often yields surprising results…principles that agree in the realm from

which we originally abstract them can, when pushed beyond that realm, produce

incompatible surprises [original italics]."[DeVidi 2012.] [40]

Sambin points out that this possibility (i.e., of having different principles

governing the same subject matter or phenomena) is actually inherent in pinning

---

[39] This notion is closely tied to the complaint that intuitionistic mathematics is merely more informative than classical math, but not a true rival. I point out how Sambin responds to this worry in the next section.

[40] In fact, Sambin demonstrates that taking his approach results in developing genuinely new mathematical structures, unavailable with other starting points [pg 86-9 Sambin 2010] (thus, intuitionistic mathematics is not merely offering a more informative method).

constructivism to a *process* of abstraction and formalization of the resulting principles.

He writes [Sambin 2010]:

> Only by "forgetting" some information can one obtain some abstract concepts
> [sic], treat them "mathematically," that is with no mention of the reality they
> come from, and finally apply them back to reality successfully. This is
> probably taken for granted by most. However, one tends to overlook the fact
> that there is no single and necessary method of doing it. Actually, different
> foundations can be seen, under this perspective, as different choices as to
> what should or should not be considered as relevant, that is, as different
> principles to be used to abstract from reality and idealize it (Pg. 67).

At this point, we need to look in more detail regarding the question of how *choice of a starting point* should be constrained if it is to be correct on constructive[41] grounds and otherwise. What we need, I argue, is a distinction among the different bases of abstraction, i.e., how much and what we "forget" when formulating principles. The distinction I have in mind is between when 'features' are included or excluded merely in order to get some desired outcome and when such inclusion or exclusion is taken (or seen) as relevant for (and as still) being *faithful* to the source notion, reality, or phenomena. The first sort bears some resemblance to what can be called a *merely*

---

[41] There is some irony in how Sambin [2010] argues for the possibility of different principles and thus different foundations. He writes in the paragraphs following the passage quoted in the previous footnote that "[i]n intuitionistic logic the notion of truth is linked with proof, and necessity, rather than with consistency, or possibility [i.e. the absence of contradiction, as is the case with classical logic]" In some sense, it seems that his argument for pluralist foundations is not constructive. This turns out to be a problem (i.e., what grounds different foundations) in other parts of his argument, such as a lamentable, misguided analogy with religious beliefs at the end of his paper.

instrumental theory, while the second has a much better claim to the title of a *constructive theory.*

Sambin and Valenti [1988] argue for something quite similar: "our principle is that an abstraction is constructive, that is, a reliable tool in getting knowledge that is faithful to reality, not when information is kept as much as possible, but when it is 'forgotten' in such a way that it can be restored at will in any moment. This after all is the test to show that an abstraction does not lead astray from reality, that is, that it preserves truth" (pg. 4).

On the other hand, if we do not care to be *faithful* to our target phenomenon, data or 'reality' we can readily admit 'features' into our process in order to generate principles that get some sort of desired results (including matching data). In the constructive case, since we must be able to *guarantee* that anything "forgotten" by principles of a more abstract level can be restored, we can prove that the more abstract principle is faithful to the more concrete level. But how can we know when the inclusion or exclusion of some feature is constructively *relevant*? The answer, I will now argue, is due to a generalization from this discussion: what I call *abstraction norms*.

First off, an abstraction norm presupposes that some but not all bases of abstraction may be *justified* in some context: some choices as to what is relevant may be justified, while others rejected (or merely unjustified). What is at stake when we so justify these choices (of a basis of abstraction) is explained by appeal to abstraction norms. In other words, abstraction is justified like any other philosophical activity: as

subsumed under relevant norms.[42] And we have already looked at one sort of abstraction norm, namely a constructive abstraction norm. So what we are looking for is not merely that we are using the abstraction for some purpose that is relevant, but whether we can justify such an abstraction as being faithful to a more concrete level *and* relevant for said purpose, all else being equal.

Thus, an abstraction norm governs not only what details will get in the way of (or just be superfluous for) some theoretical purpose but also at what degree in the abstraction process *alterations* become *distortions*. For example, and as already argued, at some levels of abstraction we might have a change in the subject matter and thus could be comparing apples with oranges. Consequently, criteria for what counts as the subject matter, 'reality', or the phenomena at hand must figure in with such abstraction norms.

At this point, I am in good shape to show what we have gained so far in my pursuit of a *prima facie* case for a plurality of foundations (that is, if you start from a constructivist foundation). Specifically, I can answer the two questions I pointed out above.

First: what makes a principle, method, or foundation *constructive*? A constructive principle, method, or foundation is the Hansel and Gretel of theoretical perspectives: it will always leave a trail of breadcrumbs so we can back track to the source. Of course, this needs for most interesting subject matters to be hedged with 'in principle,' just as the defenders of verification theories of meaning for empirical claims needed to insert 'in

---

[42] I have in mind here Herbert Feigl's [1968] argument. However, I would be remiss (in a project that explores the use of a preference logic) to not point out that there is no such thing as "pure syntax, pure semantics and pure pragmatics" even at the level of pragmatic justification, i.e., justifying a *preference* for the adoption of some mode of justification or primitive norms.

principle' clauses to avoid being committed to the claim that very distant stars never had planets if nobody would ever observe them, and various other qualifications will no doubt be required for domains that differ from the mathematical ones. But the principle of 'no in principle unknowable truths' is basic.

Secondly, what is the connection between constructive foundations and interesting, non-trivial disagreement? The connection can be found in how we can justify different bases of abstraction, that each nonetheless satisfies our constructive norms. For example, the two notions of a set, i.e., a subset is a set or not, we looked at are both constructive, as long as we don't try to have them both at once ... but they each correspond to a further goal for the abstracting mathematicians. The 'as rules' folk are motivated as describe above. The 'input output behavior' mathematicians are motivated by the idea that wherever possible *mathematics should be extensional*: that the point of mathematics is to investigate structures that can be instantiated in different ways, and that to consider two things that share the same structure as different simply because they are differently described is to betray a non-mathematical motivation. So one can qualify as a constructivist, but one can have different motivations and goals within the constructivist framework and so abstract in different ways with different resulting foundations. And assuming that neither of these is an *incorrect* motivation, we get alternative constructivist foundations. And, supposing DeVidi is on to something, we see how they can come to disagree when we extend their use beyond the 'reality' from which they were first abstracted.

Thus, to suppose that merely requiring a constructive framework in any domain will pin us down to a single correct way to abstract is unmotivated. Constructivist

motivations will seldom be the *only* motives involved in abstracting. Supposing that more than one collection of other legitimate motives is compatible with constructivism for a given domain, alternative foundations are possible.

What Sambin seems to have asserted as fact, i.e., that there isn't a single and necessary way of 'forgetting' what doesn't matter in some context, we can now see that there are in fact good reasons why such a monist posit is unjustified. This implies that it is *prima facie* possible that there is always more than one correct, constructive basis for abstracting what matters when formulating relevant principles. And we know that the results of formalizing or situating such principles in larger theories often end up being incompatible in an interesting and non-trivial manner.

We now have, at least what I hope, looks like a typical or exemplary case (i.e., in the mathematical case of the notion of *function* or *subset*) of a notion for which distinct, perfectly acceptable (and all constructive) accounts can arise because different acts, i.e., bases, of abstraction underlie them. And this typical case may also point us to a stronger argument for pluralism (given constructivist starting points). As mentioned above, what I propose is an argument with the stronger conclusion that accepting constructivist foundations *requires* accepting pluralism about foundations, is available if an additional premise is acceptable. The premise in question is that given any notion apt for abstraction, there are always be at least two distinct bases for abstractions that give rise to distinct foundations within the same general abstraction norm. That is, when we set about abstracting, and our abstraction is governed by some norm ("it must be constructive," or "it must be a liberal theory,") that norm by itself will not be so constraining that what else

is abstracted away or included cannot vary depending on the other interests the abstracters bring to the exercise.

Specifically I propose that there are always, or at least generally, supplementary interests that can be brought to bear that can be described in terms of two general types of norms: that the principles one should use, i.e., abstract out, may be characterized as either intentional or extensional (but not both if constructive). The idea here is that any constructive approach must allow for both levels of analysis, i.e. abstract, extensional reasoning and computational, intensional algorithms, but not both at the same time. I must profess ignorance on how to defend such a premise, rather I am inclined to only fear it is true (since the burden of proof returns to the pluralist constructivist, rather than the monist constructivist).

1.5 Two Rival Norms in Constructive Liberalism

Ignoring for now the stronger premise, if what I have said above is true, there should be a parallel story for constructivist liberal theories to the one just told about constructive mathematics. And in the same way, this will amount to a *prima facie* case for pluralism in liberal foundations.

In the mathematical case, constructive motivations alone can't determine an approach to answering foundational questions. Instead, for at least some foundational questions, there is a multiplicity of mathematical norms that are each consistent with constructivism—indeed, that are plausible parts of a constructive answer to a foundational question—but which are jointly incompatible when constructivism is assumed. I have considered two such norms above, namely the norms of extensionality

and of computability. I will briefly present an example of political/moral norms that might play a similar role for constructivist liberalism.

An important foundational principle in any political theory where fairness or justice is central, and so for all liberal theories, will be: *equal concerns should be treated the same*. As is probably well known, there are many difficulties spelling out exactly what this principle means; it's far more complicated than merely being a claim that decisions be principled or consistent and so constructivism alone is insufficient to determine what it means. To start, let's break down this principle into two subsidiary issues. First, what does it mean for two considerations to be equal in moral or political stature? Secondly, what does it mean to then treat such equal considerations as the same? (E.g., the world being the way it is may allow for us to act on one but not the other in identical fashion).

The obvious way to understand the second issue is that political/moral decisions must be *impartial*: whatever you do for one concern, do for any equal concern, *ceterius paribus*. This isn't to say that there isn't a lot of disagreement regarding how to theorize this impartiality. For example, Harsanyi's [1975] utilitarian notion of impartiality means treating everyone's preferences as if one super-person had them all. The best known variant of Rawls [1971] notion of impartiality is spelled out in terms of the mechanism of the *veil of ignorance;* with reasons being suitably impartial when accepted there. Scanlon's[1998] notion of impartiality means using reasons that no reasonable person would reject. Sen's notion of impartiality [2009] is quite similar to Scanlon in many ways; he argues for what he calls an open impartiality where reasoned scrutiny can come from any source, so as to help undermine parochialism.

Likely, there is an interesting story to be told here regarding different theoretical or abstraction norms at work in these various accounts of impartiality. However, I want to focus on the first subsidiary issue, due to what turns out to be an interesting parallel with the previous discussion. Sorting what it means for two concerns to be equal in moral or political stature leads to interesting questions regarding the nature of categorization or class membership, and logically minded folk find such questions illuminating. I will focus on two ways of answering this question. These two answers will, individually, be compatible with political liberalism; but under the assumption of constructive liberalism are instead jointly incompatible. Furthermore, I argue that each answer is predicated upon a different norm for what matters in making political/moral decisions.

The first answer is that when two groups (individuals can be groups) have the same to gain or lose in the outcome of some decision, the concerns of each group are equal. The norm I propose as the basis for this answer is that political/moral decisions should be focused on people's wellbeing, as befitting a focus on what one has to gain or lose.[43] As such, I will call this the *wellbeing* norm; its important feature is that it only sorts interests in terms of the *outcomes* of a decision. The second answer is that when two groups invest or entrust the same sort of resource in a decision, i.e., to the decision-makers, the concerns of each group are equal. The norm here is that political/moral decisions should be *responsible*.

---

[43] Normally, I would take such a straightforward privileging of people's well-being over other sentient animals as morally unmotivated and possibly immoral. In this case, since we are looking at political decisions, and I am assuming that it is humankind alone that is the relevant kind of political animals. This assumption might be mistaken.

To illustrate the wellbeing norm, Peter Singer's preference utilitarianism is helpful (even if one thinks his view isn't constructivist, the norm is more general than just as a part of constructivism). He is explicit in answering the foundational question we are looking at in Singer [2011]. He writes: "In accepting that ethical judgments must be made from a universal point of view, I am accepting that my own needs, wants and desires cannot, simply because they are my preferences, count more than the wants, needs and desires of anyone else. Thus, my very natural concern that my own [preferences] be looked after must, when I think ethically, be extended to the preferences of others." (pgs 11-12).

With this wellbeing approach, political decision-makers must count groups with the same preference regarding an outcome as having equal weight when making a decision. Obviously, not all wellbeing minded moral theorists are preference utilitarians, but Singer's explicit focus on the outcomes of a decision is exemplary of this approach. And if one needs an example of a constructivist utilitarian, Harsanyi can be seen as footing this bill with his version of the veil of ignorance.

On the other hand, the responsibility norm prescribes focusing on the roles people may have in the process of a decision or in the life of the decision-maker. Thus, equality of such roles in a decision implies equality of concerns, all else being equal. On this account, sorting concerns as equal requires taking into account differing concern for the resources or trusts invested in the first place. A key virtue for responsibility-based accounts is that they can straightforwardly handle cases where social roles seem to play an essential role. For example, two drowning children have an equal concern in the outcome of surviving. But if one child is your child while the other is a stranger to you,

there is something very wrong in treating their concerns [in being saved by you] as equal. Your child trusts you in a way the stranger does not. This is not to say that *wellbeing* accounts can't deal with this issue somehow, but rather *responsibility* accounts can be applied in a relatively simple and straightforward way to do so.[44] In the political sphere, *responsibility* accounts straightforwardly handle questions regarding why a preference for one's own citizens over the needs of others, all things considered, seems morally justified.

The question now at hand is whether each norm, i.e., the wellbeing and the responsibility norms, are individually compatible with political liberalism and specifically constructive liberalism. To show compatibility, I will give an example of each norm at work in liberal theories that eschew appeal to some mind-independent reality for what makes moral/political claims true.

The wellbeing norm can be seen as the basis for Rawls's approach in *A Theory of Justice*, as his principles for a just society clearly privileges liberty and then focuses on the benefits the worst-off will receive when implementing some fundamental policy. And of course, the social contract reasoning at the heart of *A Theory of Justice* is constructivist.

An explicit and simple example of a liberal theory based on the responsibility norm is a bit harder to come by. To avoid having to justify interpreting a pre-existing liberal theory as responsibility-based, I will cook one up, so to speak, to do the job. In Richard Freeman's [1984] stakeholder analysis for ethical business decisions, the moral principle at its heart is that one should maximize value for all stakeholders. And whether or not this principle is wellbeing or responsibility based depends then on how carefully

---

[44] On the other hand, an easy way to theorize responsible decisions is in terms of preferring the concerns of those who start with more to invest. And if this easy way is taken, it's not hard to see how responsibility norms could allow for principles that privilege the wealthy in a society.

we define 'stakeholder'. To make explicit how the responsibility norm can work in this approach, I will deviate from Freeman's [2002] most widely known definition.[45] Let's define *Stakeholder* as:

> Given a decision-maker *O*, a *stakeholder S* for *O* is a group whose concerns are or have been placed under *O's* care, given *O's* function(s), purpose(s) or social role(s).

This definition allows for children to be stakeholders to their parents, where there is no explicit choice on the part of children to place themselves under their parents' care. It also allows for bondholders to be a stakeholder for the organization in which they have their creditor's stake, where here there is an explicit choice to place their coupon under the organization's care. Of course, a range of examples can't directly justify a definition, but they can illustrate how the definition can be put to work.

To ensure that a political version of this stakeholder theory can be both liberal and constructivist, we need an example where individual citizens' liberty is a privileged concern placed under the government's care, but justified in virtue of a constructive method. I propose that we could construct conditions under which individuals would risk their liberty under the care of political representatives using a framework that resembles Philip Pettit's [2005] civicism. With such a framework we could avoid any ontological essentialism when it comes to delineating when a group of individuals constitute a political people. Unfortunately, such interesting details are beyond the scope of this

---

[45] Which reads [Freeman 2002], "The narrow definition includes those groups who are vital to the survival and success of the corporation. The wide definition includes any group or individual who can affect, or is affected by, the achievement of a [organization's] purposes." (pg. 47).

project. However, plausibility and compatibility should travel hand in hand well enough for now.[46]

Nonetheless, we still need to answer why it's possible to *construct* a situation where individuals will reasonably place their liberty under their governments' care.[47] Since it's possible to use the veil of ignorance to characterize an ideal contract, such that, at the minimum, we agree to a society where each has the most liberty compatible with like liberty for all; this much can be said regarding the political stakeholder approach as well. Of course, this assumes that other conditions of such an ideal contract are independent of this privileging of liberty; but this assumption must be made anyways since, as I will argue in Chapter Two, we will need a plurality of such ideal states in order to theorize liberalism—more on this in due course.

The next issue is to sort wellbeing and responsibility norms as incompatible, that is, once we assume constructivism. It is not hard to see wellbeing accounts as often and directly conflicting with responsibility accounts. To illustrate, I see the wellbeing/responsibility divide roughly tracking Nozicks's [1974] patterned/historical accounts of justice. What someone has invested (whether it's trust or capital) in some project is a matter of historical record; outcomes of some project can be articulated solely in terms of an abstract pattern. If Nozick's arguments regarding the incompatibility of patterned and historical accounts of justice are cogent, we should also expect that the wellbeing and responsibility norms are, in fact, incompatible once we assume constructivism. Here is a brief argument to this effect: reference to some historical

---

[46] That is, as long as being good constructivists, we distinguish between plausibility and truth.

[47] The simple answer: if something is actual, it's possible.

investment etc. in some project regarding what its outcome should be is clearly intensional, while wellbeing is straightforwardly reducible to questions of membership in the same group,[48] i.e., could be read as extensional, and thus logically, constructivism is undermined if we combine both of these further norms (as is in the case with constructive mathematics and combining extensional with computable norms).

But suppose we don't want to rely on the cogency of Nozick's arguments. If we can show that these different norms justify conflicting answers to basic policy questions, we would have a good grasp of what the right sort of incompatibility would look like. To do so, let me explain how I see these norms as operating at a basic level. I propose that the norms regarding wellbeing and responsibility, as I have called them, arise from the context of moral criticism in actual moral practice.[49]

When one is on the hot seat, so to speak, the charge one is accountable for, i.e. what we are responding to, is: whose concerns did you take into account in your decision? More concretely, if someone did wrong, the charge may be glossed as: why did you not think of so and so's concerns, wellbeing, rights, and/or trust etc. when you did such and such? (How could you?) On this account, principles governing the notion of *wellbeing* and *responsibility* are formulated from and with details abstracted from real-life calls for considering all the people (or even sentient beings) and their legitimate

---

[48] Of course, part of the open questions brought up here include the issue of whether there is a way to develop the notion of wellbeing in intensional terms. If this can be worked out, the incompatibility conjectured here is wrong (at least as presented for now).

[49] As you will see, I will often appeal to this context (i.e., of being called on to justify one's choices) to help specify the notions required for a suitable meta-ethics.

concerns one should in making appropriate choices.[50] The question here is: why couldn't

we formulate constructive principles that treat concerns with trust, investments etc. and

with wellbeing as on par? As in the case of ready-to-eat principles for apples, where we

seeming could combine ripeness-based and sweetness-based principles, can we always

combine wellbeing with responsibility norms in a constructive principle? To see why not,

consider the following case of moral criticism.

Suppose we are on the hot seat. We, as state officials, used taxes to help a

company outsource jobs from a local plant to a factory in a developing nation.[51] One can

readily imagine that the loudest criticism will sound something like: how could you use

our money to take away our jobs and give them to people in ….? Here, there is a claim

that we must defend a decision to use resources from one group to benefit a *different*

group. Now, instead of answering the criticism, and in the fine political tradition of

criticizing one's critics, we could query: how could you critics think families here are

more important than families there, especially since helping families in a developing

nation can result in far more benefit from the same tax dollar? This isn't to say that

responsibility based theorists won't prescribe, say, foreign aid. They may do so, but likely

in a way that by providing such aid, there will be spin-off benefits for the donor country's

taxpayers (and probably international business stakeholders). In contrast, wellbeing based

theorists would likely instead prescribe aid where the benefit to recipients is the greatest

---

[50] This foundation is largely motivated by Bernard Williams' [2012] focus on taking moral phenomena, especially personal and social responsibilities (not to be confused with my technical notion of a norm of *responsibility*), as the starting point of analysis (as opposed to a systematic or theoretic approach) regarding what changes to the world one should bring about.

[51] 'Why?' one might ask. Perhaps we were moved by the plight of those suffering from a catastrophic variability of economic activity, and thus privileged those under conditions of absolute poverty relative to our own relatively affluent citizens.

overall. And of course, it's a contingent fact which country would benefit the most from aid and which country, provided with aid, would provide the most spin-off benefits for the donor's taxpayers. Thus we would have interesting, non-trivial, constructive but conflicting answers to fundamental policy questions about obligations regarding foreign aid.

If I am right regarding the nature of pluralism in constructive mathematics, there should be a parallel story for constructivist liberal theories. I argued that, as in the case with math, constructivism alone couldn't determine how to answer some fundamental policy questions in liberal theory; we need to apply further abstraction norms to provide the relevant answers. This will also result in a *prima facie* case for pluralism in liberal foundations. I argued that sorting which concerns are equal when making political decisions can be based on two constructive abstraction norms: *wellbeing* and *responsibility*. For the former, political decisions should be focused, i.e., given a privileging of liberty, on the wellbeing of those involved. For the latter norm, political decisions should be responsible to those who have entrusted their care to the decision-makers. I gave an example of two constructive liberal theories, each based on each norm, to show that the norms are individually compatible with constructive liberalism. Then, I argued that there are two ways in which these norms are instead jointly incompatible once we assume constructivism. First, the wellbeing norm is extensional in nature and the responsibility norm is intensional in nature. As with the case for mathematics, constructivism can't be combined with both extensional and intensional accounts of group membership. The second argument showed that each norm gives rise to conflicting principles for directing foreign aid, and thus disagree at a fundamental level.

1.5.2 *A Plurality of Prescriptions*

On this account of the nature or foundations of constructive reasoning, it is *prima facie* the case that there are multiple correct, but disagreeing, ideals from which we can abstract out prescriptions, obligations, principles of justice etc. As mentioned, these prescriptions etc. are not merely claims made by the relevant ideal; rather if the ideal is cogent, what it prescribes is in fact prescribed. Thus we will have some inconsistent prescriptions in such cases; to make pluralism intelligible, we need to know how to make sense of such a possibility. And this is part of the contribution this project is aimed at providing.

1.6 *Conclusion*

Summarizing so far, I have looked at why, if we want a systematic and constructivist meta-ethics for liberal foundations, we will need to allow for such a system to be self-inconsistent. And to investigate such possibilities, it would be appropriate to use formal tools. Specifically, I have argued that Bernard Williams' and Amartya Sen's claims, i.e., that taking pluralism seriously is tantamount to rejecting systematic constructivism in ethics, leaves open at least one avenue of escape. And this is the option I propose: that we should allow for a constructive system to be inconsistent. I briefly and informally sketched out how the sort of pluralist approach I propose, i.e., that allowing for the system to be inconsistent, we are able to respond to their objections.

But this talk of theoretical inconsistency is likely hard to handle. The idea is that while we may want to use a *pluralist perspective* to defend a foundation for political liberalism, we first need to justify a method for investigating the details of a suitable

76

pluralism. I argued that the relevant moral terms and concepts can be treated as particular sorts of mathematical objects. In particular, I introduced and informally sketched out the technical notion of a *choice*. Given an appropriate formal language to reason with and about such *choices*, we will be well on our way to understanding the ramifications and consequences of a pluralist approach for liberal foundations.

Finally, I responded to doubts regarding the epistemic status of pluralism in meta-ethics. It may have been hoped that the multiplicity of moral approaches to liberalism is merely an artifact of the state of current scholarship, *contra* Williams or Sen. However, I argued that the very nature of constructivist foundations, when properly understood, generates a *prima facie* case for a plurality of such foundations. It would appear that a constructivist approach leads one to allow for inconsistent systems, in general and in meta-ethics in particular.

Chapter Two   *A Feasible Plurality*

2.0 Introduction

In Chapter One I suggested that a proper foundation for liberalism ought to be a *pluralist* one. In this chapter, I begin the defense of this claim.

I will begin by considering how best to understand the phenomena of moral conflict and disagreement in order to clarify what a correct pluralist account should look like. I then turn to the important task of defending pluralism against some obvious objections. This part of the chapter is not merely defensive, though, since the shape a defensible pluralism must take has important implications for meta-ethics. As it turns out, taking conflicting principles seriously requires a nuanced and complex conception of disagreement and incompatibility between rival theories.

2.1 CA and Value Conflicts

CA methodologies seem (but only seem) to have a better grasp on how to handle conflicting principles than RITs. As Sen [2009] pointed out, RITs seemingly suffer from what he calls the infeasibility problem when faced with conflicting principles. Since CAs specify which principles are socially preferable (i.e., which principle is selected from some specified impartial perspective) they promise to resolve disputes in any particular context. Consider two principles, say *A* and *B*. And suppose that in some context, *c*, *A* prescribes some choice that *B* prohibits. Perhaps, *B* always prohibits lying but *A* says you should not put someone's life at risk without good reason (all other things being equal). The context is a person having a heart-attack and she asks you if she is having a heart-attack. Since you have basic CPR training you know that staying silent or telling the truth

will likely worsen the symptoms or cause death. Given a social choice methodology, we could input principles *A* and *B* together with assumptions about how each option will ramify in distributing 'goods' across society. Presumably, in this example, a reasonable CA will opt for the principle that allows us to tell 'white lies' in life-and-death circumstances. More generally, though, one option will be better or at least no worse than the other option in how well it distributes goods according to the preferences of our social perspective.

However, using CAs does seem to imply that the principle of *we should always maximize the preferences of the impartial perspective* will never be over-ridden by any other principle, on pain of circularity or nonsense. In other words, upon closer examination CA methods do in fact presuppose a uniquely correct principle. Since for such CA methodologies, *justice* just is a species of choices preferred by impartial perspectives,[52] we must give up the possibility of conflicting principles at a fundamental level. What may have looked promising, i.e., in terms of taking conflicting principles seriously, in ordinary moral contexts turns out to be unsuitable for handling the more foundational issues at hand.

I am assuming that there is not just one way of appropriately thinking about justice, and Sen claims to make a similar assumption.[53] I agree with Sen that our store of

---

[52] Non-cognitivist responses to the open-question argument in some sense relegate all questions of action-guidance to individual motivations to act. In this one aspect, Sen's [2009] appeal to actual preferences could be seen as immune to the open-question argument. However, Sen's belief that we can then appropriately aggregate such motivations (preferences) in a theoretical structure (i.e., the impartial perspective) revives the argument.

[53] Pgs 203-207 Sen [2009] argues for other foundations of political reasonableness than the notions of reciprocity and mutual benefit as found in the social contract tradition.

such impartial reasons and notions regarding justice do seem to outstrip our ability to capture every relevant consideration in one unique *and consistent* set of principles. However, for any relevant CA all these conflicting principles of justice can be attributed to different accounts of what claims can be defended as appropriately impartial.[54] The problem is that even if all reasonable considerations are given due weight and concern within any and every suitably impartial perspective, this does not imply that there are still no other reasonable ways to think about justice. In fact, this seems to suggest another open-question: But is it reasonable to only defend claims from impartial perspectives?[55]

It looks like taking *justice* to be a species of choices preferred by an impartial perspective doesn't allow for us to represent all the conflict we do find in actual moral practice. If this is the case, CA methodologies flounder at a fundamental level, given a pluralist starting point. On the other hand, it is clear that CA methodologies can take conflicting principles in more ordinary contexts quite seriously. And as you will see, I will borrow ruthlessly from what makes CA methods work at an ordinary level to help RIT approaches handle more fundamental disagreement.

2.2 "Ought" Implies "Can" and Monism

There is another common view regarding the nature of intractable conflict between principles: there are none. I will briefly and informally introduce this view by using an argument proposed by Terrance McConnell [1976;1978]. McConnell shows that the conjunction of the following three sentences express an inconsistent set of

---

[54] See footnote 53
[55] And as suggested in Chapter 1, there are some reasons to think that a *responsibility*-based perspective is possible.

propositions: (1) there are genuine moral dilemmas; (2) *ought* implies *can*; (3) if one is

obligated to do each of two courses of action, then one is obligated to do both courses of

action. McConnell does not offer a formal specification of to his deontic language, but it

is not hard to identify the reasoning behind his argument. In keeping with normal practice

in deontic logics, **O**A stands for "it is obligatory that A" and ◊ stands for classical alethic

possibility. Premises introduced to represent the above propositions will be noted by [1],

[2], or [3] to keep distinct from rules referring to steps in the argument.

1) **O**A   premise

2) **O**B   premise

3) ~ ◊ (A & B)   premise; 1-3 represent [1]

4) **O**(A & B) → ◊ (A & B)   premise; by way of [2]

5) (**O**A & **O**B) → **O**(A& B)   premise; by way of [3]

6) **O**A & **O**B   1,2 &I

7) **O**(A&B)   5, 6 →E

8) ~ **O** (A&B)   3,4 m.t.

Since 7 and 8 are a contradiction, some step leading to them must go. McConnell's

choice is line 3, and so he puts the fault on the assumption that there are genuine moral

dilemmas. He makes this because the formulas of #4 and #5 are axioms in his version of

standard deontic logic (SDL). But this is too fast; why shouldn't we instead revise the

axioms or logic being used to represent deontic deductions? There are technical reasons

for rejecting a polymodal logic combining *KD* and *S5* (the closest well-understood logic

to approximate McConnell's take on SDL). For example, it seems like the naturalistic

fallacy (that is, a formula reasonably interpreted as such) is satisfied by all models for such a polymodal logic (details can be found in appendix 1).

Less technically, it seems problematic to use a formal representation to rule out one side in a live debate, unless that formal representation is taken to be indisputable by all relevant sides. McConnell seems instead to have the tail wagging the dog in this argument. As I will show, the inference from *ought* to *can* is not in general valid if we take the distinction between *ought* and *must* seriously, let alone the possibility of moral conflicts. Moreover, the formalization of obligation that McConnell seems to have in mind is not constructive, however the details of such an argument is part of a much larger project.

In Chapter Four I developed a logic for characterizing the deontic consequence relation that rejects both [2] and [3]. At worst, then, even if we assume McConnell's line of reasoning is correct, we have a stalemate since the axioms of my deontic logic permits us to posit the existence of moral dilemmas if we so desire. However, as I will show, there are independent reasons for preferring my system.

Given the sort of considerations raised by Sen, it is safe to assume that there are many, diverse impartial reasons that actually exists for thinking of justice in one way or another. Our store of such impartial reasons and notions regarding justice do seem to outstrip our ability to capture every relevant consideration in one unique *and consistent* set of principles. If at least some obligations follow from our notions of *justice* then a logic that is at least neutral with respect to the existence of moral dilemmas is far more appropriate than the 'standard' logic assumed by McConnell and those who share this view on this issue.

2.3 On a Family of Liberalisms

Indeed, it is the recognition of this failure of any single, consistent theory of justice to capture every relevant consideration that has motivated other philosophers to consider some form of pluralism. For example, the many counter-examples offered in the literature to Rawls' principles of justice from *A Theory of Justice* tend to show how there are legitimate instances of injustice that Rawls's principles fail to pick out [Copp 1974; Keyt 1974; G.A. Cohen 1991; Harsanyi 1975]. As noted in the introductory chapter, near the end of his life Rawls explicitly argues that his principles of justice should have perhaps only a minor role in justifying public policy. Instead, he argues there should be a family of liberal theories that make up the Public Reason that may be used to justify policy. Rawls' move to a plurality of RITs was motivated by this recognition of the many, diverse and seemingly correct reasons for thinking about equality and liberty that actually exist in our sort of democratic culture.

Let us pause here to distinguish the case for pluralism from the mere recognition that people sometimes disagree. So far I have relied on an intuitive sense of what sort of disagreement matters in a philosophically worthwhile pluralism. To develop a more nuanced notion of pluralist disagreement, I think a useful place to start is David DeVidi's [2012] argument for mathematical pluralism. His claim is that for pluralism to be a philosophical view worthy of philosophical attention, it must be, when spelled out, *interesting* (e.g. it can't turn out that we've all been pluralists all along) and *nontrivial* (the differing theorists cannot be mere variants, sub-theories of the same overarching theory, etc. – they need to disagree!)

While DeVidi had mathematical and logical theories in mind, this seems a plausible requirement for moral or political pluralism as well. And it is a condition I think Rawls' Public Reason meets. There are, of course, subsidiary conditions pluralism must meet—the different theories of the same thing—so that the disagreements are not merely argument, involving a subtle changing the subject from one theory to another and so on.

We can develop this pluralist notion of disagreement further: the sort of disagreement that we are looking for should have rival theories or ways of seeing the same subject matter. But not just any rival theories will do; for the disagreement to be philosophically worthwhile, the rivalry must be intractable. In this case, *intractable rivalry* must mean two things: 1) at a fundamental level there is some sense in which the rivals agree, e.g. on what metaphorical race they are running; 2) that the criteria for success is non-excludable,[56] i.e., that it doesn't follow from a theory satisfying the criteria that no other theory can.

It should be clearer now why the recognition of rival theories or abstract principles could motivate notions similar to PRIT. In fact, when we look at the theoretic resources PRIT has to handle the infeasibility objection, these resources are due in large part to the structural features PRIT shares with other interesting and non-trivial pluralisms found in other domains.

In the math case, a non-trivial and interesting pluralism implies that there may be more than one correct solution to a well-posed mathematical problem. To be clear,

---

[56] It is plausible that this might mean that inquiry in mind-independent or realist domains cannot then be pluralists in this sense. A plausible criterion for a successful theory in such domains is: which theory gets the world as it really is? And unless the *world* has parts that can be represented by true contradictions, it seems that there are no rival but equally successful theories.

pluralism on this account needs *multiple correct* answers to a well-posed question.[57] As argued in Chapter One, given a constructivist starting point for liberalism, there is a *prima facie* case for pluralism. Given these considerations, such pluralism implies that the bases for such rival theories or formal principles are used to abstract said principles etc. from the same concrete level. On this account, rivals must be faithful to the same more concrete notion. If this is the case, there can be no aspect of the concrete notion that can decisively rule out all but one of the rivals. Since the infeasibility objection is what it is because of the plurality of impartial reasons, if we were to posit rival theories, rival ideals, it is possible to then account for such a plurality.

Recall that PRIT is the view that there is at least two equally (and maximally) correct but distinct accounts, i.e. rivals, of what a perfectly just institution consists in. (While, like later Rawls seems to suggest, I argue there are more than just two such rivals, and all we need for pluralism to hold is that there are at least two.) PRIT is therefore similar to other interesting and non-trivial pluralisms from other domains: to be an interesting and non-trivial sort of pluralism the view must be that a particular phenomenon, domain, or subject matter can be *correctly* specified or explained in more than one *genuinely distinct* manner. The shorthand version for these criteria in the case of PRIT is that the cogency of PRIT rests on the claim that there is more than one correct answer to a well-posed question regarding the nature of ideally just societies.

In the political context, I will use Rawls' organizing question (but choose your own if you want): How should we balance the values of *liberty* and *equality* in the basic

---

[57] To be clear, it is a theoretical pluralism, not just methodological pluralism. Methodological pluralism is more often defended on the grounds of the incompleteness of a method leading us to use other methods to fill this gap.

framework of social cooperation? Suppose we can justify different starting points in how we specify these values or what counts as impartial or what counts as a good balance. The result will most likely be a non-trivial and interesting pluralism. In the liberal RIT tradition, I explicate a particular balance of these values by trying to describe a perfectly just society based on those starting points.

For PRIT, the many, diverse impartial reasons that actually exist for thinking different things unjust are understood as parts of rival accounts of a perfectly just society. To be explicit, PRIT is an interesting and non-trivial pluralism since we have but one subject matter, namely *what is the nature of a perfectly just society?* And I propose that, for the purposes of the present inquiry, we can assume we have at least two genuinely distinct but correct answers. Each correct account of a perfectly just society can be seen as formulated upon a distinct group of the many, diverse impartial reasons for thinking different situations unjust. There can be some overlap in these groupings of reasons; however, we assume that there must be at least one difference when we assume the pluralism to be non-trivial.

As pointed out above, I agree with Sen that our desire to capture every relevant consideration in one unique *and consistent* set of principles of justice is a false hope. But I contend that there is good reason to suppose that for any one of the many, diverse, impartial reasons that could be a relevant consideration in theorizing justice, we have good reason to think it to be included in one of the *plurality* of accounts of a perfectly just society. Here's why: if you consider what an impartial reason would have to be like in order for it to be independent of every regulatory ideal, such theory-independence will not look plausible. Such a reason could not:

1. Play any role in modeling the decisions of a just society

2. Be entailed by any correct version of an ideally just society.

3. Be a premise in the construction of any theory of ideal justice.

4. Be added to an ideal theory to construct an improved theory (that is, extending a theory if it is consistent to do so).

5. Approximate the sorts of reasons used in policy decisions in any correct version of a just society.

The idea here is it seems implausible that such impartial reasons could not count as touchstones for a theory of an ideally just society. Conditions 1-5 survey the different ways rivals can be formulated upon these relevant impartial reasons or imply such reasons are acceptable. 1 and 3 are clear examples of how rivals may take aspects abstracted from the more concrete level and then formulate principles based on these considerations. Conditions 2 and 5 point out the 'discovery' of a new sort of reason we should think of as relevant or impartial, either in formal terms or in approximate form as the case may be. And condition 4 indicates how we might be able to accommodate principles in the context of pre-existing theory. Each condition captures how such reasons can count as touchstones for our various rivals, either as criteria in their formation or as tests of their results.

It seems unlikely that any significant consideration could not be accounted for as a touchstone for at least one rival version of an ideally just society. Obviously, this argument is not conclusive; it is not intended to be so but rather it is conductive in nature.

2.4 Pluralism and Paraconsistency

An obvious objection looms, so let me pause to address it. The cost of a PRIT framework is that deductive reasoning from a PRIT perspective must be paraconsistent. Assuming that we have an interesting and non-trivial pluralism there must be some p such that one theory permits p while the other does not permit p. As I have developed the notion, our family of liberal principles of justice must be jointly inconsistent in order to be an interesting and non-trivial sort of pluralism. They must disagree and that disagreement cannot be explained away by each theory dealing with a different subject matter or some other form of equivocation. To be clear, the individual theories themselves need not be inconsistent; inconsistency only arises when we combine them into a family of rival theories. But if we assume that the pronouncements of each correct theory about what is obligatory and what is not, then if deductive reasoning is to be fruitful under these conditions, we need a logic that does not verify the rule of explosion.

Could we avoid this conclusion? Why not change the properties of the theories in our family of RITs so that such inconsistency does not arise automatically? Perhaps we should give up the claims that each acceptable theory be complete. That is, should we allow individual RITs to have truth gaps (i.e. claim lacking both truth and falsity) instead of truth gluts (i.e., claims being both true and false)? In this way, we could have disagreement between the members of our family of RITs that amounts to one theory permitting p and a different theory failing to say one way or the other whether p is permissible; thus, there would be no need for paraconsistency, at least at face value.

Two problems arise. The first problem is that such disagreement is not vigorous enough to be non-trivial. In this domain it seems always justifiable to appeal to the

stronger theory in such cases, even when the theories are rivals in other cases. In other words, if one theory is silent on some issue but a rival rules it just or unjust, the silent theory does not guide action in the way its rival does. This seems to suggest that being silent is a theoretical vice in such a case; i.e., the action-guiding theory has a *prima facie* claim to being a better theory. This would render such disagreement trivial.

Moreover, there is an additional aspect to this problem. Supervalutation techniques allow us to fill in truth gaps of incomplete theories. Supervaluation treats the incompleteness of a theory in terms of its underdevelopment. Thus, given assumptions about how things will turn out theoretically, truth gaps are only relative to a stage of theoretical achievement, progress or available evidence, information etc. With such formal representations available, there would have to be independent justification for not using filled-in-theories instead of the gap-py theories. That is, we could always in theory generate a stronger theory than an incomplete one.

The second problem questions the conceptual plausibility of a theory of an ideally just society failing to be maximally complete. It is true that our actual theories in use in contemporary public debate are not maximally complete, let alone generally correct. However, RITs are not merely cleaned up versions of actual beliefs held by citizens; they are theories that specify an ideally just society. It is not plausible that an ideally just society would be less than maximally complete for the following reasons. First, there is a conceptual connection, albeit an inchoate one, between an ideal society and complete set of action-guiding principles at the heart of the RIT approach. The idea here is that there is something sub-ideal with a society that doesn't know which policies to implement. There is a distinctly comparative or attributive element to this notion of an *ideal society*: My

society answers more policy questions than yours so mine is better, i.e., the more comprehensive theory, consistent with the same question is better than a less comprehensive theory. In an important sense, the current project can be seen as an attempt to make explicit this inchoate connection between an ideal society and action-guiding principles.

Secondly, in this sort of normative research we are aiming at providing action-guiding principles for collective choosing, everything else being equal; this goal seems to incline us to construct maximal theories instead of less maximal theories. It is only when we find discrepancies between equally correct maximal theories that we bemoan such completeness. We have a conflict here between the theoretical values of consistency and action-guiding/maximal completeness. I have noted already the difficulties with living up to our action-guiding duties in a pluralistic theoretical framework. We should not choose options to make this harder than it must be. One quick clarification: by maximal or complete theories in this case I do not mean a theory that is formulated upon every impartial reason. Recall the distinction between infeasible incompleteness and theoretical incompleteness from the discussion of Sen's infeasibility argument. In this case, the completeness is in terms of ruling for any P whether P is true or false. The above argument for an ideally just society providing a complete theory is by no means conclusive. But until such justification is provided, the balance of considerations decisively tilts to using the truth-glut approach instead of the truth-gap approach.

2.4.1 PRIT and True Contradictions

There is a meta-ethical and metaphysical worry I want to head off. If prescriptions are truth-apt, and mixed inferences[58] suggest at least some are [Tappolet 1997; 2000], and PRIT is true then it seems plausible at first blush that some dialethias (true contradictions) exist. For instance, suppose we have some policy, *P* and two rival *liberalisms X* and *Y* that disagree about the justness of *P*. If both *X* and *Y liberalisms* are equally applicable for evaluating *P* then it seems that *P* is both just and not just. It seems we are contradicting ourselves, if we are not equivocating on the meaning of justice.[59] We could bite the bullet and accept one of the options, i.e., that we are truly contradicting ourselves or equivocating on 'justness' (but in an interesting and non-trivial way). However, neither option seems all that appealing at first pass. True contradictions are unappealing for obvious reasons. Equivocal use of *justice* would seem to undermine the notion of being rival accounts, i.e. in terms of the interesting disagreement criterion, of an ideally just society. Nonetheless, I shall adopt versions of both these options. The plan is that careful and nuanced formulations of both these options *together* will provide a sensible account of what is going on when rivals disagree (as required by a philosophically worthwhile pluralism).

On the first option, dialetheism in PRIT may not be as strange or troubling as perhaps first thought. In the next chapter I will develop tools for dealing with trumping

---

[58] The mixed inferences I have in mind are arguments with both descriptive and prescriptive premises. For example: Killing John is wrong (prescriptive). Mark is killing John (descriptive). Therefore, Mark is doing wrong by killing John. The pattern of inference is valid according to the standard account, so truth-preserving. The premises must be truth-apt, or so this account implies.

[59] This is a modification of Graham Priest's [2006; 2008] argument against logical pluralism.

reasoning in ethics. In doing so, I will show that the best way to construct formulas

representing prescriptions will be in terms of a prescribed choice.[60] Thus it is not that we

have a proposition and its negand both being true, as is in the case of the standard

presentation of dialetheism. Rather I posit that we have two jointly incompatible choices

both being *acceptable* or "true"—more on this in the next chapter.[61]

2.4.2 Equivocating on Justness at different levels of analysis

But there is a further problem regarding incompatible properties, such as *justness*,

I want to highlight. Dealing with this problem will lead us to see how there can be a sort

of equivocation on justice on one level of analysis that still allows for the relevant

theories to be rivals on a different level of analysis. The idea is that by representing

different sorts of incompatibility we can allow for an interesting and nuanced

equivocation rather than merely changing the subject matter. Moreover, by having

different *incompatibilities* we can provide a useful perspective for regimenting the

relation between moral properties, prescribing choices and moral theories. The problem I

have in mind is a sort of over-determination issue, in that actions may have properties

independent of any theory's assertion (contrary to a constructivist account of the

attribution of properties).

---

[60] More formally, a prescription will be a relation over tuples of actions and sets of
such tuples modified by a deontic box or diamond (with negation when required).
[61] If something like this view is correct, together with the notion of mixed inference,
it implies that *acceptable* and *truth* are just species of a more general notion of what
is preserved in valid inferences. Given my predisposition towards type-theoretic
foundations seen in earlier chapters, this should not be all that surprising. But what
this more general notion is exactly is beyond the scope of this project.

As I will argue in Chapter Three and Four, it is choices that are, in the first instance, subject to moral evaluation; actions have properties only due to their role as options from which a choice must be made. But for this to get off the ground, two actions compared in this way must not only be rankable as better or worse, but also must have a measure of how much better one is that the other.[62] If such comparisons can measure how much better an option is than another, i.e. a cardinal ranking, then the following problem for pluralism can seemingly occur.

Suppose we interpret the acceptability of a choice, according to a theory, as implying that the selected option is *more just* than the rejected options. Given that a theory is comparing actions to other actions (and assuming that a theory may be in some cases indifferent between trading off a certain number of actions against another type of action) the *justness* of actions can be cardinally ranked by a theory. Assume we have two theories *A, B* and three actions *x, y, z* but that *x* and *y* are incomparable according to *A* and *B*. Suppose that *A* states that we should choose *x* over *z* but not *y* over *z* and *B* states that we should choose *y* over *z* but not *x* over *z*. The problem is that if theory *A* states that *x* is much more just than *z* but *B* states that *y* is barely more just than *z* then (if we are not equivocating on 'justice') *x* is more just than *y* is relative to *z*, even though we have no theory stating that *x* is more just than *y*.[63]

---

[62] A comparison of actions is cardinal when it also compares lotteries composed of subsets of those actions, i.e., when an indifference curve can be constructed. Thus a cardinal ranking will tell us how much better one option is than another.

[63] Of course, as stated this is a problem for my project only in principle. After all, I have already argued that the relevant moral or political codes or theories should be complete. But the problem could be modified to deal with decision-procedures for handling conflict between such codes—more on this in Chapter Three.

It seems that what started out as an interesting and non-trivial pluralism regarding ideally just societies threatens to turn into monism. Here's why. If moral theories cardinally rank options and we can unequivocally attribute *justness* accordingly then we seem to have a sort of inter-theoretical comparison[64] that can over-rule any given theory regarding what properties obtain.

I want to approach this over-determination problem by finding a way to equivocate on *justness* as applied to specific choices while treating the *justness* of an ideal state unequivocally. In other words, we want a principled reason to move between a sort of theoretical relativism when dealing with claims regarding the justness of specific choices and a theoretical pluralism when dealing with claims regarding what we ought to do, in ideally just states or otherwise. By theoretical relativism, I mean that predicates are indexed to the theory used: given a predicate A and theories Y^n, if X according to theory Y^n is A_n, we have relativism. Framed this way, theoretical pluralism is: given a predicate A and theories Y^n, if X according to theory Y^n is A, we have pluralism. On this view, theoretical pluralism abstracts away or ignores the source, i.e. the relevant theory, of the predication. So the question now is, what abstraction norm can underwrite this base of abstraction?

We could first find out when we shouldn't abstract away the source of predication. For example, we already have a rule that specifies that the justness of choices (or any other property) is a matter of theoretical relativism (and because such a rule wouldn't allow the ranking of x, y and z in the manner described above). But it's not a very philosophically interesting rule: given the fact that we are committed to constructive

---

[64] Compare with how interpersonal comparisons of utility is used in social welfare economics to aggregate individual's preferences, given a suitable ordinal ranking.

norms, theoretical relativism lets us exclude the faulty sort of ranking. Obviously, this

seems more than slightly *ad hoc*; I think we can reverse-engineer a more principled rule.

If a multiplicity of theories in one case calls for pluralism, and we know the bases for this

requirement, we might have a better view for understanding when a multiplicity of

theories calls instead for relativism.

The idea I have in mind is that *obligation* can be, and should be, treated as a

logical word (as argued in Chapter One), but not so with moral properties. We may use

moral property-claims, which obtain relative to some theory, to justify an obligation; but

justifying an obligation by using a moral theory doesn't imply that the resulting

obligation inherits any theoretical connection to that theory, given that it is a logical word

(understood in terms of a suitable deontic logic). Once we have an acceptable obligation,

we kick away the theoretical ladder, so to speak, that let us defend it in the first place.

Logical words don't care about content, i.e., the particular theory used, that is, once they

are recognized as (deontic) logical in nature. Therefore, when representing obligation in a

formal system we must ignore what theory was used in deriving said obligation since

obligation is a logical word.[65]

Thus, the rule I have in mind is that when dealing with a multiplicity of theories:

if the subject matter, i.e., the concepts, predicates etc., can be and should be treated as a

logical word, pluralism is needed; otherwise, relativism is called for. That is, given a

---

[65] While I maintain no commitment to Rawls' notion of overlapping consensus, I think there is some happy coincidence here. On his account, where the reasons used in reaching some consensus come from don't matter. As long as everyone involved can use their own views to defend the target claims, it shouldn't matter what the reasons are, all things considered. Fortunate for my account, I am making use of a far weaker set of theoretical assumptions than does Rawls, especially as my arguments generally rest on a logical analysis of what the relevant terms can mean.

multiplicity of correct theories, pluralism abstracts away which theory was used, as called for when the subject matter is a logical word; but not so with relativism: each claim is indexed to the theory used. Obviously, there are important details to be worked out here; however, defending such a pluralist basis of abstraction is a far larger project (especially, how it relates to the arguments in chapter one regarding constructivism implying a *prima facie* case for pluralism). Nonetheless, something similar to this basis of abstraction is at work throughout Chapter Four and some details can be found there.

2.5 Different Pluralisms

A final question still remains regarding pluralism. Why not instead deny that equality of correctness implies equality of applicability? I will call this option, "Jurisdictional Pluralism". On this alternative, there is a distinction between the sorts of answers equally correct theories may cogently disagree upon. The idea is that evaluative questions are kept distinct from theoretical questions. For a more widely understood example (but one where I want to bracket off the issue of whether it is merely a methodological pluralism or a fundamental one), say we have two models for representing fundamental physics. The first understands a basic element as a wave and the second represents the same thing as a particle. Clearly the obvious ways to interpret these models disagree on questions on the nature of this element of physics. However, given different experimental setups, the context will indicate that one of these theories is inappropriate or appropriate for predicting experimental outcomes.

My notion of jurisdictional pluralism claims that non-trivial and interesting disagreement may take place only at a descriptive or theoretical level; no disagreement

occurs at the evaluative or predictive level. The idea is that the nature of evaluative or predictive questions relies in large part on the *purposes* we are aiming for in making this assessment. On the other hand, non-evaluative questions are asking for a description or abstract representation of some phenomenon. A PRIT version of jurisdictional pluralism would then claim that only one *liberalism* can ever be applicable in evaluating whether a particular policy is just or not, even though other *liberalisms* are no less correct in how they describe the nature of ideally just institutions.

In this project, I assume that the less problematic option is to let incompatible choices be jointly acceptable. I think we can make sense of the notion that policies can be both just and unjust: acceptable in one ideally just society but objectionable in a different ideally just society. On the other hand, there is no clear rule for when one of our rival liberalism would be exclusively applicable for evaluating a particular policy, as required by the jurisdictional account. Of course, an argument from ignorance is no argument at all. Rather, I propose that both sorts of projects should be of some value, but only the one is in play in the current work.

2.6 Conclusion

I started off with looking at two monist ways of looking at conflicting principles or obligations. In particular, I showed that 'ought' can't imply 'can' if we allow for conflicting obligations, i.e., genuine moral dilemmas. And if we use an impartial perspective to decide all questions of conflicting principles, we rule out reasonably disagreeing about using an impartial perspective. Finding fault with both these monist frameworks, I showed that a philosophically useful pluralism needs *multiple correct*

answers to the same well-posed question. I argued that we should understand such an interesting and non-trivial pluralism as needing *rival* theories. Rivals disagree in a particular fashion: Rivals must be complete regarding the question at hand.

I then looked at some plausible objections to theoretical pluralism, including whether pluralism is a sort of dialetheism and questions of when we are mistakenly equivocating on the relevant terms. I argued that the semantic values underlying the notion of *obligation* and its cognates should be understood in terms of *acceptable* or not. If I am correct, a true contradiction of obligations isn't as bizarre or worrying as if obligations were truth-apt like, say propositions about apples on tables. Next, I argued that we could formulate the distinction between relativism and pluralism in terms of pluralism ignoring the theory used in predication, whereas relativism indexes such claims. With this formulation in hand, we could then look to whether the relevant terms should be treated as logical in nature. This technique should then answer whether a term should have an unequivocal sense or whether terms should be indexed to the relevant theories.

Finally, I briefly looked at an alternative way of understanding pluralism, what I called 'Jurisdictional'. But without a clear sense of the norms we could use for uniquely applying different liberal theories to particular evaluative questions, I suggested we stick with the account already at hand. And as it stands, we have a good enough view of theoretical pluralism that we can now turn to looking at how the notion of *obligation* must adapt to such a context.

Chapter Three　　　　　　*Ordering Obligations*

3.0 Introduction

In this chapter I begin the work on how we should adapt the notion of obligation (and cognates), if we are to fruitfully use it in a pluralist context. First, I will be concerned with the problem of action-guidance in a pluralist context. The idea is that, when faced with the sort of disagreement seen as problematic for action-guidance, the relevant codes may prescribe trumping the original but conflicting obligation with a coordination mechanism. These mechanisms could allow us to settle disagreement in some, if not most, cases in a liberal society. Clearly, there are two subordinate issues involved in this strategy. First, what is a trumping relation? Second, what is a fair coordination mechanism?

First up, I look at the philosophical problems with formalizing trumping-relations in deontic logic. This part of the project allows me to resolve two problems. First, using such a formal perspective offers a powerful method to respond to a large part of Sen's redundancy objection to RIT foundations. Second, and more helpful at a general level, this perspective shows how the action-guiding problem raised by disagreeing principles can be made far more reasonable, and manageable from a practical point of view. But the reasons for looking at these issues go beyond the immediate demands of this project. Trumping-relations are ubiquitous in moral, political, and legal reasoning and so deserve to be taken seriously in any deontic logic. But trumping-relations are largely ignored in most deontic logics. This isn't surprising since *trumping* poses many inconvenient problems for those constructing formal systems.

3.1 The Persistent problem with trumping

The different consequence relations formalized in logical systems are for the most part what we call monotonic relations. This means that the premise set from which a specified conclusion set is deduced can be freely extended with additional premises.[66] For example, if *A entails B* then *A and C* entail *B* for any *C* whatsoever. Monotonic logics include classical and intuitionistic logics, as well as normal modal logics. However, trumps-reasoning is not monotonic: the purpose of trumps reasoning is to change what is permitted or obligatory, as the case may be, by introducing a new premise. Relevant logic is not monotonic in the sense defined above.

However, trumps-reasoning is also not symmetrical. For example, *not letting someone die* may trump *not telling a lie* but *not telling a lie* rarely trumps *not letting someone die*. In relevant logics the order of premises does not matter to what may be deduced. However, with the non-symmetrical nature of trumps-reasoning, the order in which premises are introduced is significant. The usual method for handling the difficulties posed by this sort of consideration is to eliminate rules that let us ignore the order of premises, which results in a more complicated semantics for negation. Essentially, a non-symmetrical negation of incompatibility would need to be added to or replace the standard negation of being false. I will not follow the usual method. Instead, I will introduce a technique that instead simplifies trumps-reasoning. In fact, I will not introduce any extra semantics for negation to deal with trumping relations. But to be fair, this choice is purely pragmatic; the change from the familiar logic we learned in

---

[66] Characterized by the structural rule of weakening: $A \vdash C \Leftarrow A; B \vdash C$—see chapter four for details.

undergrad to relevant and constructive versions is extreme enough. Further deviations from the familiar are likely just unhelpful at this point.

I will first show how the tools available in current non-standard deontic logics fail to accurately represent trumps reasoning. This will allow me to explain how my approach differs in a way that productively addresses these problems. I then take this technique and solve two problems. I will show how members of a RIT plurality can use coordination devices they prescribe to solve some, but only some, disagreement between members in our family of liberal theories. This result will mitigate some of the action-guiding worry mentioned in the introduction. But perhaps most importantly for current purposes, these coordination devices are used to solve concrete policy problems. In the light of these results, the redundancy complaint loses much of its bite.

The difficulties existing approaches have with trumps reasoning are due to the *persistence* of truth. By the persistence of truth I mean that if a proposition is true in our semantics, that proposition will remain true regardless of further developments. The persistence of truth is normally, and I assume correctly, regarded as a conceptual truth. Recall that my starting point is that moral necessity=obligation, and with *standard* approaches this means that moral necessity is represented by the truth of a proposition at all morally accessible worlds. If we assume that trumps-reasoning with such a propositional moral necessity requires a non-symmetrical negation, (so that the order of premises matters to which premise trumps which premise), the persistence of truth is impaired.

The problem with this framework is that if an obligation can be trumped by a different obligation, truth at a moral world is no longer persistent (or as we shall

sometimes also say, *hereditary*). Persistence would fail if a proposition, say *A*, true at all morally accessible worlds could be rendered no longer true at all morally accessible worlds because of trumping. To see this possibility, assume that *OA* is true in the actual world, i.e., *A* is true at all morally accessible worlds, but that *OA* is trumped by a different obligation, say *OB*. The truth of the trumping obligation, *OB*, requires that there be a proposition true at all accessible worlds that is incompatible with *A*. But *B* and *A* cannot both be true at all morally accessible worlds, if there are any. In other words, if *OB* trumps *OA*, it can no longer be permissible or obligatory to make *A* true. Thus, the semantics would have to somehow record that *A* would be true at all morally accessible worlds, if it were not for the fact that *OB* trumps *OA*. We started off thinking that we could represent *obligation* in terms of truth at all morally accessible worlds! However we now have to represent the notion, at least in part, in terms of a conditional logic or counterfactual model of some sort. In other words, given this approach, we must treat the trumping relation along the lines of: if it weren't (or isn't) the case that *OB*, it would be the case that *OA*.

While I grant that there might be a case to be made for using conditional logics to represent trumping, I think such an approach would miss out on an important fact about trumping-reasoning. The fact I have in mind is that a trumping relation is also its own (or introduces a new) obligation: an obligation regarding the *order* in which you may do certain actions. In an important sense, if *OB* trumps *OA*, you are obligated to make *B* true *instead* of *A*. Here's what I mean: suppose we assume *OA* is true at the actual world and that *OB* trumps *OA.* If we then also assume *OB* we are not just obligated to make *B* true but also to not make *A* true after all. This fact suggests that in order to represent

trumping-relations, we should do so in terms of an obligation to order actions in some particular fashion. If this is true, representing obligation as *propositions* true at all morally accessible is wrong-headed. In a nutshell, this is why you will see me use order-theoretic tools and what is called neighborhood semantics for representing moral necessity in Chapter Four, instead of the more familiar universal quantification over accessible 'worlds'.

3.2 Why not 'Obligation' as suitably modified actions?

As a first step towards representing trumping relations as an obligation ordering actions, I begin with the hypothesis that, instead of modifying propositions with deontic boxes and diamonds, they should modify a logic of actions. Let me explain: suppose we have a sentence expressing an action and modify it with a deontic box or diamond. In other words, it is actions (not propositions *simpliciter*) that are prescribed, proscribed, or permitted as the case may be. Important for current purposes, the order of actions matters in a properly constrained logic of action (as we will shortly see in more detail). Thus, this approach seems a promising start for constructing sentences that express moral obligations and permissions, in particular, obligations to do actions in a particular order.

Using a deontic logic of actions rather that propositions may require other adjustments that may be problematic for use in deontic logic. The rules for combining *actions* should not include commutative rules, but standard logics involving propositions do include such commutative rules. (An example of a commutative rule is that the order of premises doesn't matter.) In other words, I can combine some actions in one order but not in the reverse order. Consider an example, due to Greg Restall, where I can go buy a

car and then give all my money away but I can't first give away all my money and then buy a car (all other things being equal). One can fine-tune one's logic using commutative rules of differing strength, i.e., with different types of reordering being acceptable. It might look, given the issues raised about trumping and how order matters in what obligations trump which others, that removing commutative rules is a good fit. But as I have already mentioned, there are significant costs when we turn to representing negation, and thus necessity, and possibility.

The first issue is that losing commutativity seems to introduce very fine distinctions between different senses of obligation and its cognates into our semantics that don't generally arise in actual practice. Here's how: when the order of premises matters, we must have two implication arrows instead of one in our logic. The reason for this, while technical, is not hard to see. When we combine two premises to prove some conclusion, the deduction theorem states that we can also infer that one of the premises can prove a conditional with the other premise as antecedent and the original conclusion as the consequent. That is from $A; B \vdash C$ both $A \vdash B \rightarrow C$ and $B \vdash A \rightarrow C$ follow. However, the deduction theorem must be refined when the order of the premises matters, as two different kinds of conditionals follow from $A; B \vdash C$. The first uses the standard left to right arrow that we learned in intro logic: $A \vdash B \rightarrow C$. The second sort of arrow records that the order of the premises in the original proof matters; it is a right to left (or backwards) arrow: $B \vdash C \leftarrow A$. It is not easy to gloss this distinction in ordinary language, but perhaps we could say for the first sort: from A we can deduce that if we add B then C. The second can be read as: from B we can deduce that provided that we had A in the first place then C. The distinction between 'if we add' and 'provided that' disappears when

order doesn't matter; but in the current case 'if' represents adding information and 'provided that' represents a prior assumption. As we will see in the next chapter, in this system negation of a formula is defined intensionally (i.e., by reference to 'worlds' other than the one it is 'true' at). We can use arrows to represent this definition of negation, which implies that there would be two sorts of negation (i.e., right to left and left to right), and so two ways of representing being obligated *not* to do something (or not obligated as the case may be) depending on which sort of negation is in play. This is probably an example of how a subtle distinction in an artificial language sometimes doesn't matter for understanding or representing the corresponding ordinary concept. So, this distinction is of value for deontic logicians, and probably very few others. In other words, positing different senses of *prohibition* etc. in order to interpret non-commutative rules in a formal system may be seen as too cumbersome to be worth the trek. If we could do without adding these extra versions of obligation, all the better.

3.3 Order puzzles and actions

A connected, but more philosophical, problem is that the nature of actions in of itself should not dictate moral theory. But this could happen in improperly constrained deontic action logic. A good example of this can be found in the various *order puzzles* common to the literature. While the details on how my proposal resolves this class of problems is part of a larger project, the basic issue with such puzzles resolves around specifying a rule for the order in which obligations are to be discharged. (For example, representing trumping-relations is a type of order puzzle.) The nature of action combination is in many ways at the heart of order puzzles. Since actions are

105

asymmetrically incompatible, we have a puzzle with how to specify the moral accessibility relation so that we couldn't represent the permissibility of arranging our actions in such a way to avoid doing something we think we, pre-theoretically, must do, and so solely in virtue of the nature of action combination. The problem with so specifying the moral accessibility relation is that we seem to need to refer to the particular *content* of the relevant worlds, which can't be done formally, i.e., *a priori* or in axiomatic terms.

For example, to be responsible for some harm, a person should be mentally competent. But getting drunk renders one temporarily mentally incompetent. At first glance, it looks like getting drunk and thus temporarily mentally incompetent absolves you of responsibility for harms you inflict after you become drunk. Indeed, it is a (I think sad) fact of legal and moral history (from 1949-1982 in U.S.A.) that showing that the accused was drunk was successfully used as a defense in sexual assault cases in just this way [Hasse 1972]. Improperly constrained, the logic can justify past judges' decisions to find accused rapists not guilty because being drunk rendered them mentally incompetent. As mentioned above, being able to appropriately specify such a constraint on the logic is deeply problematic. This isn't to say that there might not be many promising attempts to do so. My proposal is that most of the problem can be dealt with in a far simpler way— but more on this in the next section.

One final example before we move on. The issue I have in mind comes from the idea that '*ought* implies *can*'. This claim seems to imply that morality is silent when there are no options. Suppose now that a person performed actions at an earlier time that have the effect that at a later time she has no other options but to, for example, maximize profit

within the bounds of the law. Then, on this account, moral obligations are no longer operative. Why should the order in which actions can be combined have such ramification for moral theory? We need more than an appeal to an ambiguous platitude as is '*ought* implies *can*'. I think we can do so.

3.4 Obligation in terms of Contrastive Principles

While the technical details are part of a different project, my philosophical response to order puzzles such as representing trumping relations is to use what I call the *choice* structure. The basic idea is that representing the relative moral ordering of actions should use a complex structure of the relevant actions, rather than what actions are true in accessible worlds.

To motivate this technique philosophically, I will introduce and develop what I call the concept of *contrastive principles*. In a nutshell, we are quantifying over ideal answers to a contrastive "May/Ought I" question when we are looking for action-guidance. A contrastive "May I?" question is regimented by the *choices of actions over alternatives*. In other words, the question queries the moral nature of choosing one option from a specified group of alternative options. Understanding principles in terms of answers to contrastive "May I?" questions (that is *choices* instead of actions *simpliciter*) will permit me, or so I will argue, to avoid deontic problems with the order in which actions are performed.

First, a quick tangent regarding what is at stake when representing *actions*. Of course, whether or not the actions in the system are clustered into *choices* in some way, it remains true that *actions* are being represented. So it may well seem that I owe an

account of what I take actions to be. It seems to me, though, that the only theoretic

assumption I need to make about actions is that there can be multiple tokens of the same

type of action. I do not deny that different tokens of a single action will have different

empirical descriptions. However, if we can treat such actions as equivalent for the

purposes of developing intentional explanations then a type-token structure is appropriate.

On this account, actions are individuated in terms of the role they play in intentional

explanations. However, PRIT is neutral regarding the nature of intentional explanations.

3.4.1 Contrastive Principles

Before I can explain and justify *contrastive* principles, it is important to try to

clarify how *principles* figure into this system. I take it as a starting point that principles

guide our actions, either implicitly or explicitly. Moreover principles are what give

content to moral theories that specify the Right. Since RITs specialize in specifying what

an ideally just society consists in and thus a version of the Right, we are looking for a

plurality of principles, given RIT pluralism. I will call 'choice-principles' those abstract

principles we are willing to commit ourselves to employing when deliberating about

deciding future concrete cases. I propose that a choice-principle is, to a first

approximation, the answer to a question, "May I do X?". However, decisions, as even

elementary decision theory correctly teaches us are about trade-offs. If decisions are

about selecting one option from among alternatives, there is something conceptually

amiss if we take the answer to "May I do X?" as fully capturing what matters in a choice-

principle.

Considering actual choices, it is clear that any choice is drawn from among a class of alternatives. If the class of possible alternatives picks out the important aspects of the decision at hand, we will want in our theory a way to represent the distinction between "May I do X instead of Y?" and "May I do X instead of Z? The contrast class is the set of alternatives (Y in the first case and Z in the second) to be considered as available options we could choose instead of the one actually taken, i.e., *X*. Adapting from Van Fraassen [1980], I will call the question's targeted option the *topic*, in this case X. For example consider the topic of lying. "May I lie instead of telling the truth and people die as a result?" questions the permissibility of lying, the topic, differently than does "May I lie instead of telling the truth and lose 2 million dollars as a result?" Since I am modeling principles as answers to *May I* questions, I will also refer to a principle's *topic* as well; the answer is supposed to tell us whether we can choose the topic instead of a member of the contrast class.

Analogously to the literature on explanation, I call this refinement of permissibility norms '*contrastive principle*'. To a first approximation, I define a *moral principle* as a universal (morally speaking) answer to a *choice* question, i.e. "May I do X instead of Y?" And as expected, *permissibility* is a negated universal negation but in my system, this does not imply that *permissibility* can be represented as an existential quantifier. Briefly, here's why: having one *yes* answer to a choice question is not equivalent to it not being the case you ought not to choose so, since we are taking moral dilemmas seriously. But more on dealing with having both *yes* and *no* answers to a *May I?* question later.

Employing *contrastive principles* gives this theory important flexibility. For instance, it makes it possible to model various subtly different sorts of *May I?* questions. As we shall see below, being able to model the change in questions allows us to solve a lot of different problems, including order puzzles.

The first and most obvious way of changing a *May I?* question is by changing the composition of the contrast class. To repeat the example that I raised earlier, we could change the consequences of telling the truth from losing two million dollars to someone dying. While both *May I?* questions regarding the topic of '*lying*' seem to have a token of the type '*telling the truth*' in the contrast class, the consequences of acting in such a way are quite different. And we can structure the components of a contrast class to reflect this complexity. The truth is, that the contrast class differs for each *May I?* question in terms of a different type of action. In the first case, it will have a token of the type '*if you tell the truth then you will lose two million dollars*'; in the second case, the token is of the type '*if you tell the truth then the patient will die*'.

A second way of changing a *choice* question is by changing the *contextual factors* relevant to giving a satisfactory answer. The idea here is that the question may be framed by conversational clues, contextual constraints, or outright specification for a particular type of answer. For example, say you are an astute listener [and presumably are talking to a philosopher] and you pick up on a conversational implicature that the person wants a Kantian answer. You gather from this implicit demand that the question requires more than a simple yes or no, but requires a specific sort of explanation as to why Kant would have said yes or no, under ideal circumstances. In other words we need to model how we can preserve the *satisfaction* with an answer.

To tie this in with the previous account of constructive abstraction, if I want to safely use the notion of *moral principle* in what follows, it is best if I show how we can recover what is 'forgotten' in the abstraction process. As in the mathematical case, it looks like we can formulate theoretical principles governing *May I?* contextual factors either extensionally or intensionally. For instance, we can formulate a theory of contextual factors intensionally, by listing or forming a set[67] of variables representing contextual concerns that can be 'checked off' when we evaluate an answer to a *May I?* question. On the other hand, we could formulate a theory of contextual factors extensionally by directly treating elements in the model as denoting the type of contexts salient to *May I?* questions.

I argue the right approach is to choose an intensional method. Most of my confidence in this approach comes from a deeply ingrained prejudice for computable types of constructivism. However, if you don't share this motivational framework, more can be said. Here's what I mean: if we treat "because" as part of the statement form of answers to *May I?* questions, the resulting structure cannot guarantee preservation of satisfaction over substitution of co-extensive expressions. The statement form (of a positive answer) I have in mind is the following.

> *Yes, you can choose (topic) instead of (contrast) because (topic) is (blank) but (contrast) is (blank)*.

For example, consider the possibility that the topic of some choice is both the greatest good for the greatest number and can be willed as a universal law. Moreover, suppose each member of the contrast class is equally good or bad in respect to each moral theory,

---

[67] In this context, fuzzy sets or lists are not off the table. We might need degrees of membership to get the right results.

i.e., for any *y* of the contrast class, both theories claim that *y* has the same moral property relative to the topic. I am well aware that this supposition is completely false for the real world versions of these theories. However the issue here is logical, not whether I am being faithful to these theories; if you want a more believable case, some Kantian theories may be very close to having the same extensions as other Rights-based accounts. Suppose now that the *May I?* question is being asked by a stalwart defender of rights and as such finds any appeal to treating the good of individuals in terms of some composite super-person deeply unsatisfactory. In this case, even though choosing the topic belongs in the extension of both moral categories and each element of the contrast class has the same value regardless of the moral theory used, only one of these two categories may be part of a satisfactory answer for this stalwart Kantian. But this means that the categories of contextual features cannot be extensional after all since on supposition all the terms are co-extensive (i.e., $\forall x(fx = gx) \rightarrow f = g$), that is, for any *x* of the contrast class). In other words, our imaginary Kantian wants to know *how* you came about in deciding that the topic is an acceptable choice over the alternatives.

Modeling this piece of language/behavior under this condition (i.e., of the failure of substitution of co-extensive terms to preserve satisfaction) indicates that an intensional approach is more appropriate. As in other cases, such as *it is possible that, it is known that, it is believed that,* an intensional approach seems best served by constraining how terms in an extensional semantics are *related* to each other. For instance, in *possible world* semantics, this is accomplished by using the accessibility relation. By configuring this accessibility relation in different ways, we can develop different extensional semantics for various intensional operators.

In the *May I?* case we are looking for an answer that says something along the lines of "with this information together with such and such moral theory you can reach, get to, access, favor, or accept the topic but not the other contrast class members." To be more precise, the structure of a minimally satisfactory answer looks more like: "no other member of the contrast class is more reachable, accessible, favorable etc. from the combination of said information and moral theory than is the topic." This more precise version jettisons any assumption that there is a limit to the comparative intervals of favorability, etc. It also abandons any assumption that there is among the options one unique option that is most favorable, reachable, etc.[68]

In contrastive explanation, contextual factors are modeled by using a *relevance* relation that an appropriate answer must satisfy between the question's topic and its contrast class. Borrowing terminology from the literature on *contrastive explanation* I will say that a *May I?* question is a triple $Q = \langle T, R, C \rangle$, such that any appropriate answer to Q must offer relevant information and theoretical considerations R, as to why the topic, T, may or may not be chosen instead of other members of the contrast class C.

To further develop this account, it will help to look more closely at how choice questions are answered. I will start with a brief and informal typology of answers. Consider the *choice* question "May I lie to make 2 million dollars instead of telling the truth that my product is more risky than previously thought?"

1. Yes.

2. You can lie instead of informing the public of the risks because the market will set the price for any increase in risk.

---

[68] Of course, this amendment is in many respects analogous to David Lewis's [1973] revision of the Stalnaker conditional.

3. You can lie instead of informing the public of the risks because consumers have no right to information they don't pay for.

4. No.

5. You can't lie instead of informing the public of the risks because such risks have more marginal costs for the most vulnerable in a society

6. You can't lie instead of informing the public of the risks because it is unfair since the worst-off in society would be better off if they knew the additional information.[69]

The negative and positive answers are clearly parallel, so I'll just discuss one: let us look at the negative answers. Number 6 is a *direct moral* answer as it explicitly uses a piece of moral/political theory to answer the question and offers no auxiliary information. Number five is a *direct contextual* answer since it assumes the moral theory of 6 and makes explicit the grounds assumed in 6, i.e., that the worst-off would be better off knowing the risks. If we combined 5 and 6, we would have a *complete moral* answer. Not complete in terms of accounting for all and any moral considerations, but complete in explicitly offering relevant information and data from the salient moral theories and from the context. Number 4 is *code* for the other negative answers. The same sort of relationships holds for the positive answers.

Perhaps the most important lesson to draw from this typology is that some answers, i.e., the direct contextual sort, can presuppose a particular moral theory or facts from a context. This indicates that a *choice* answer can be rejected (or changed) on multiple counts. And as argued above, by changing what answers would be satisfactory,

---

[69] Adapted from Van Fraassen [1980]

we thereby change the question asked. We can change the options under consideration by changing the contrast class. When we change a question by changing what we hold true in a context, we change the *relevance* relation characteristic of the question. In changing the *relevance* relation we challenge the presuppositions of the *choice* question; by changing those *direct context* answers which are satisfactory we might challenge which moral theory should be used to help decide the issue. Or we can challenge facts about the context. From the example above, we may challenge number 2 because risks external to the parties directly involved in a transaction generally fail to be accounted for in the price mechanism of the market. If the product has such externalities (the indirect costs mentioned above) then the market will not set the price for all the risks involved in selling, distributing, and using the product. In the sense indicated, the answer misunderstands what the question demanded or assumed in what would count as a good answer, in this case: the facts.

Choices can then be specified to be as finely-grained (or not) as required by the phenomenon in question. This framework provides considerable flexibility in the level of specificity of questions because both the contrast class and the relevance relation are variable. Like *contrastive explanation*, changing the contrast class or *relevance* relation changes the relevant aspects of the topic under investigation. And this allows different sorts of contrast classes or *relevance* relations to represent requiring or presupposing different sorts of moral or political theories in satisfactory answers.

No matter whether we want to specify an action intensionally, intentionally, extensionally, externally, collectively, cooperatively, individually or even divinely we can set up a contrast class or relevance relation to correspond to however we choose to

specify the target action/topic. So far it should look like anything goes in terms of what could count as a relevance relation since any direct answer to a *choice* question will bear at least one sort of relevance relation between the topic and the contrast class. Appearances are deceiving since constraints on the relevance relation are not specified *a priori* but regimented by the selection of classes of worlds where the propositions that represent answers are true. This selection of classes is what we call a neighborhood of worlds, but more on this in the following chapter.

3.4.2 Choice-Principles and Moral Sentiments

So far, I have offered what may be seen as a fairly syntactical account of moral principles. After all, in the theory I am offering, permissibility and other moral norms are understood in terms of choice-principles. And before I turn to showing how we can use choice-principle structures to represent trumping (and other order puzzles), It is appropriate to make a short foray into what I propose is a powerful philosophical interpretation of choice-principle structures.

Choice-principles, I propose, should be formulated from elements drawn from what I call justified choice-attitudes. By using the term 'justified choice-attitudes' I mean to indicate that some of our attitudes to how our choices affect others may be justified, while other attitudes unjustified. Again, the concrete level I argue we should abstract our more formal notions from is the context of moral criticism. The relevant type of criticism is less focused on what rules you followed or not as the case may be, but at an even more concrete level. The type of criticism I have in mind cuts right to the heart of how your

choices appear to reflect due, or its lack of, concern for others' well-being, rights, interests etc.

On this account, PRIT meta-ethics is a cognitivist approach, and constructivist about the role of moral sensibilities in theory. With this interpretation, PRIT may then be characterized as explaining moral judgments in terms of which moral sentiments are justified under ideal procedures. An unjust situation is one in which someone in an ideal process should find unjust. Alternatively, ideal procedures allow us to simulate what an ideal sensibility would find just or unjust. But as constructivists, we have to give reasons for why such a procedure can faithfully model ideal sensibilities. This notion assumes that sensibilities are subject to criticism, training, or degradation and that we can make explicit certain moral properties (and their structure) to explain why someone's sensibilities are better than another's.

However, even if one were to reject this interpretation for choice-principles nothing regarding the core (i.e., logical) framework of PRIT rests on this specific account. Nonetheless, this specific interpretation of the foundation shows up again later in an interpretation of the relationship between choice semantics and the truth-conditions for obligation in Chapter Four; this should be unsurprising given the terminology outlined above.

This being said, let me sum up so far, since I am about to put these structures to work. We start with a sentence expressing a choice of an action over a specified set of alternative actions. We then modify this sentence with a deontic box or diamond. All things being equal, this is the best hope for developing an artificial language whose

sentences can represent moral obligations and permissions. In other words, it is *choices* that are permitted, prescribed, or proscribed as the case may be.

3.5 Contrastive Principles and Trumping

I now return to how I account for the phenomenon of *trumps-reasoning.* We may need to represent prohibiting, for example, choosing one sort of action over choosing two actions of a different sort (saving one evil person's life or saving two innocent kittens). Principles, in essence, may need to be capable of being fine-grained.

To this end, I start with interpreting the atoms of the system in terms of choices of a *tuple*[70] of actions over a specified set of *tuples* of alternative actions. Then I modify some of these sentences with a deontic box or diamond, as the case may be. By interpreting our logical atoms as choices we have a much better way of expressing moral obligations and permissions. For example, let C be a *choice* in the sense just described. The salient tuples are constructed from a set of action-types A. For example, let $A=(a, b, c, d)$. Then let $\boldsymbol{A}$ be the set of all tuples constructed on A. We define a *choice* as $C \subseteq \boldsymbol{A} \times \boldsymbol{A}$ such that the first argument for C is a single tuple and the second argument may contain any number of tuples. We can then use each $c \in C$ to represent a sequence of action-tokens that could be chosen at the cost of not choosing other sequences of action-tokens

Consider the horrible situation of saving beloved pets from a burning house. Let a be the action-type of saving a dog, b the type of saving a cat, c the type of saving a hamster, and d the type of saving a bird. A couple of examples of possible choice

---

[70] While there are many ways of characterizing *tuples,* the method I find easiest: a tuple is a sequence of nested ordered pairs. It is essential for the use I have in mind that tuples can have multiple instances of the same element.

relations would be: 1. $c_1 = \langle b, b, a \rangle (\langle d, c, d, c \rangle, \langle d, d \rangle)$ which represents the choice of first saving two cats then a dog. The alternatives in this case are: saving a bird then going to the hamster cage and grabbing a hamster then returning to the bird cage to grab another bird and then returning to the hamster wheel for the last hamster; or just saving the two lovebirds in the same cage. 2. $c_2 = \langle c, c, c, c, a \rangle (\langle b, d \rangle, \langle a \rangle)$ which perhaps represents the choice of grabbing the entire cage of hamsters then whistling for the dog instead of chasing the cat and bird around the burning house or just saving the dog. What tuples of action-tokens are available to choose among depends on many facts about the case. These facts often lie outside of moral theory, of course—how fast you run, how obedient your dog, how hard it is to carry the birdcage etc.

With these tools, we can represent a sort of trumping-reasoning by modifying *C's relata*. Suppose *y* is a choice of $\langle a, a, a \rangle$ over $\langle b \rangle$. For ease of presentation, I am for now going to use the idea that moral necessity is a matter of truth at all morally accessible worlds. Of course, this changes in the next chapter for the reasons explained above, but nothing in what follows depends essentially on one notion or the other of moral necessity. Let *y* be true at all morally accessible worlds, i.e., in all such worlds one chooses to save three dogs over saving a single cat. This is consistent with a choice, like *x*: $\langle b, b \rangle$ is to be chosen over $\langle a, a, a \rangle$ also being true at all such worlds. We can have both *x* and *y* true at all morally accessible worlds with no inconsistency or incoherence resulting: it is obligatory to choose doing an *a*-action three times instead of doing a *b*-action once, but saving two cats is always chosen over three dogs.

Let us return to the question "May I tell a lie instead of telling the truth and having a person die as a direct result?" Suppose that our starting principle-answer, *P*, is

"It is permissible to lie if it saves lives." However, upon further investigation we learn there are good reasons to think that the young man whose life would be saved by the lie told will in turn torture hundreds of innocent kittens. Suppose that we then affirm the principle, *Q*, "It is impermissible to lie to save the life of serial kitten torturers." This example is chosen because looked at apart from the contrast classes the claims look inconsistent; one entailing the permissibility of saving the person the other says it is impermissible to save. But this inconsistency is *only* apparent. The choice-question has clearly been changed, and each principle answers a different question. So both can still be true at all morally accessible worlds with no conflict. To be clear, while the principle *Q* trumps *P* in this case, nonetheless *Q* and *P* happily coexist in our intended semantics; which principle trumps which principle is in large part up to the question asked.

To represent trumping in terms of a change in the *relevance* relation, the technical apparatus that will be developed in the next chapter is needed. Informally, like the previous account such trumping can be understood as changing what counts as a satisfactory answer to a *May I?* question.

3.6 Trumping and CTD Obligations

In many respects, this picture of trumping shares a structure with CTD obligations. Ideally, you should be able to get your way, i.e., do what your code prescribes. But in the sub-ideal conditions of disagreement between moral codes, it is often best to find some compromise. CTD structures are all about sub-ideal conditions. In the case of CTDs like the Gentle Murder Paradox [Forrester 1981], ideally you should not perform some evil, such as murder. But let's say, as the story goes, that you do go about planning on some

murder. Clearly, this is a morally sub-ideal condition. Nonetheless, it is best, if you do

kill, to do so gently. The issue that CTD obligations bring to deontic logic at large is how

to make sense of the closure of obligations under known entailment. It seems to make

sense that if *A* implies *B* and you are obligated to choose *A* that you are thereby obligated

to also choose *B*. But hold on here; if you ought to kill someone gently, killing someone

gently implies that you killed him; thus if you ought to kill someone gently, you ought to

kill him (*simpliciter*). And since conditionals are thought to be transitive, we can derive:

If you murder, you ought to murder. This puzzle can be thought to suggest we need to

index obligations to their respective conditions of moral ideality. But we need to do so

without thereby equivocating upon "ought" or its cognates. Technically, in the proof-

theory in Chapter Four we accomplish this by not letting in a structural rule called 'modal

modus ponens' and in the semantics by using neighborhood functions to determine

obligation.[71] This rule lets us combine two necessities (in our case, obligations) into one.

The philosophical issue however is how to make sense of not adding in this rule. I turn to

this issue now.

The similarity with trumping under conditions of disagreement with CTD

obligations suggests that if we have an order theory that represents reasoning about CTD

obligations, we will thereby have an order theory that can be used to reason about

trumping. Thus, the hope is that the order theory justifies not adding in modal modus

ponens.

---

[71] However, you won't find explicit reference to modal modus ponens in the proof
theory, as for ease of presentation there I approach this particular issue instead
solely through the lens of the neighborhood semantics.

However the similarity between CTD reasoning and trumping-relations can be misleading at first. This similarity may seem to suggest that we should override a more ideal principle with one less ideal. In this case, it looks like a better option is being trumped by a worse option and that is a problem. While the formalization of these concepts will provide the needed tools, some preamble may help explain what is going on. We do not move down the ideality scale, so to speak, to represent the trumping relation. Rather the disagreement or other sub-ideal conditions trigger the need to find out what should guide us given these sub-ideal conditions. Once we find ourselves at a lower level of ideality, we are looking to move up. In other words, it is the principles in terms of being involved in the disagreement or conflict that are being trumped. Thus, in an important sense we have to keep clear at what level the relevant principles can be found. And if this is the case, a rule allowing us to always combine obligations into one, such as modal modus ponens, couldn't be safely added. The formal order theory that follows will clarify this notion. Not only will we be able to get clear on how we can reason about and with CTD structures and trumping, but with the underlying order theory we will be all that much closer to defining a deontic consequence relation, given the pluralist framework of PRIT.

3.7 Trumping, Cooperation and Coordination Devices

Recall that there were two motivations for analyzing how obligations can be trumped. The previous sections dealt largely with the logical and general philosophical issues. The second issue was that I plan on using trumping relations to respond to worries that disagreement in a pluralist context renders the action-guiding power of such moral

theories void, or at least problematically muted. And now that I have contrastive principles on board I can turn to this proposal that some but not all disagreement between RITs can be solved by appeal to shared coordination devices. In so doing I need to show that acceptable RITs need to include principles that specify which coordination devices may be used and under what circumstances. Moreover, this step will render Sen's redundancy objection largely implausible. Thus a pluralist turn does in fact provide a helpful defense for an RIT approach.

After this result, I show how a class of these devices is to be mobilized when disagreement between two distinct but equally correct RITs occurs. Under such circumstances, the relevant RITs must share a principle that amounts to: when there is disagreement, and all other things are equal, using such and such coordination device trumps the original prescription within each RIT.

As it is now understood, an ideally just society specifies the terms of fair cooperation among its citizens. However, any instance of cooperation requires coordination and this coordination itself can be just or unjust, fair or unfair, perhaps based on a lottery that may be unjustified or justified. Thus each RIT, to be adequate, must contain at least one account of what counts as a fair, just, moral etc coordination device with its characterizing equilibrium selection mechanism, i.e., some procedure to decide who contributes what to making everyone the best off they could be. More precisely, by 'equilibrium' I mean distributions such that no one can improve her lot given the choices made by everyone else involved. To be clear, we could have two equilibria but that the biggest winners in one distribution are the least improved in the other and *vice versa*. The terms of fair cooperation would have to specify how to fairly

123

coordinate on one distribution instead of the other.[72] If adequate, such RITs may then appropriately model the terms of fair cooperation. For if such equilibrium selection mechanisms are just, fair and so on independent of RITs then something similar to Sen's redundancy objection will return. That is, an RIT must posit how we are to coordinate on collectively choosing one equilibrium option over another.

Appropriate RITs can then justify concrete policy decisions in virtue of the inclusion of equilibrium selection mechanisms within each characteristic Ideal. In other words, concrete policy choices can be treated as instances of coordination and RITs must have something to say about fair or just coordination. I now turn to use coordination devices within the political realm itself in order to resolve, in part, action-guiding concerns.

There are many types of cooperation (social, family, economic, military etc). The need for coordination devices is not restricted to social and economic life but is needed in the political sphere as well. I propose that an adequate version of PRIT requires each RIT to include principles that permit cooperation with those who hold to conflicting RITs.

---

[72] This is one way of explaining how early Rawls went wrong in *A Theory of Justice*. He claimed any distribution that maximally improved the lot of the worse off was fair, all other things being equal. However, this theory of fairness was lacking an equilibrium selection mechanism, since who could end up as the worst off could (and likely would) change in different 'worst-off are best off' equilibria. After all, there is no worst-off until we structure a society in a particular manner. For example, suppose we have two groups of people, the Brainy and the Brawny. Suppose there are two equilibria, each worst-off group is no better off than the worst-off in the other option. But in one option the worst-off are the Brainy and in the other option the worst-off are the Brawny. The question from any worst-off group arises, if we were to arbitrarily select one equilibria: why should we have coordinated on this option when we would have been the better off group in the other equilibria. In other words, why is it fair that we cooperated to make you the better off group when you could have cooperated to make us the better off group? As you can see, fair coordination matters just as much as fair cooperation.

More specifically, adequate RITs in RIT pluralistic worlds such as ours will include *principles* governing what to do when other acceptable RITs disagree with our favored RIT. Adequate RITs, on this account, must recognize the acceptability of other RITs in the political sphere, and in so doing must then include terms of cooperation with such disagreeable partners.[73]

Suppose we have a case where such disagreement arises and each RIT involved has a principle that asks us to use such and such a coordination device to resolve the disagreement (a coordination principle). Now if each RIT involved had the same coordination principle, the disagreement could be resolved. In this case, all relevant RITs would need to have, for whatever reason they see fit, at least one such coordination principle in common. To be frank, it's possible that an RIT may not override a conflicting prescription with a coordination principle. One way this could happen is that a stalemate may be seen as more approximately just than the chance an opposing RIT would get its way. Perhaps more disastrously, conflicting RITs may not share a coordination principle in common.

3.8 Conclusion

I first looked at why truth-apt propositions make for terrible atoms in a deontic logic that can represent trumping-style reasoning. Even when we try a conditional logic, we miss out on the fact that a trumping relation introduces a new obligation: how we

---

[73] In Chapter Four, I prove that a version of this claim follows from a weakest possible logic appropriate for deontic reasoning. You may think of this claim as a refinement of Rawls's notion of reasonable pluralism in a weakly well-ordered state, but nothing essential hangs on this parallel. Nonetheless, if this parallel is correct, it shows that we can get something like Rawls' claim but without his theoretical assumptions.

ought to order the fulfilling of obligations. This led me to try to represent trumping as an obligation ordering actions. Here we saw that the nature of action combination poses some difficult issues for representing obligation.

Instead of trying to specify how the nature of action combination and obligation can work together, I suggested we have a simpler and more illuminating option: choice-principles. I argued that choices and choice-principles far more directly capture what matters when we are looking to a moral theory to guide action. Next we saw how the structure of choice-principles models trumping as changing the question at hand. To finish off this line of argument, I briefly and informally explained how contrastive principles could solve order puzzles more generally by showing the similarity between CTD (contrary-to-duty) structures and trumping relations.

With these preliminaries set up, I turned to using my theory of trumping relations to explain how even when moral theories disagree, we still might have action guidance. Here I introduce and defend the distinction between fair coordination and fair cooperation. Fair cooperation is a matter of identifying what sort of equilibrium we may choose; fair coordination is a matter of identifying how we should choose a particular equilibrium state so identified: the equilibrium selection mechanism must itself be fair. With the idea of fair coordination we have a powerful philosophical interpretation of how and why moral theories, if they disagree, may nonetheless jointly decide upon a method to compromise (or what I call "political cooperation"). But of course, it is neither a matter of logic nor meta-ethics whether two disagreeing theories will always have a coordination mechanism in common. This means that we may still allow for intractable and principled

disagreement in a liberal society, as befitting a foundational theory trying to keep actual

practice as a touchstone.

Chapter Four   All the Different Best Options

4.0 The Order Theory for PRIT

To sum up so far: given a RIT framework, formulating a democratic and liberal notion of justice seems to require a pluralistic approach. Together with the notion of *contrastive principles* to account for the trumping phenomenon and other order puzzles, I seem to be at a good starting point for developing a formal meta-ethics for PRIT. It is also fairly clear at this point that by taking a pluralist approach with theorizing justice (including the notion of *contrastive principles*) there will be some rather drastic changes in how we should structure our concept of *obligation*. And as I have argued, using the formal tools of deontic logic looks like a promising, albeit perhaps surprising, way forward in such an investigation.

4.1 Preliminary Issues

Thus far I have argued that *choices* should replace atomic propositions as the parameters of a suitable deontic logic. I begin my formalization of *obligation* in earnest by developing some ideas that naturally arise if we take an algebraic approach to our models. In usual formulations of algebraic semantics for various logical systems, we allow for the possibility there are many more truth-values than merely 'true' and 'false,' and assume that any such truth-values are (partially) ordered. This approach lets us think of logical consequence in a particularly perspicuous manner. That is, when the relation of semantic consequence, i.e., $P \vDash Q$, holds between two propositions, this will correspond to it being the case that in every interpretation P is *less true than or equally true* to Q. In a deontic logic such an algebraic ordering can be glossed as a moral ordering since if P

being good (for example) is less true than Q being good, this relation is (on this account) equivalent to saying that the latter is morally better than or morally as good as the former. And as in the case with 'truth' (i.e., if P is true but less true than Q it follows that Q is true is well), when P is a good thing to do but less good than Q, it follows that Q is a good thing to do as well.

There are three issues that naturally come up at this juncture. The first question is fairly formal in nature.

(O&O) Given a moral ordering of choices, how should we identify which choices are obligatory?

Sure, an ordering tells us which choices are better than others but what is the connection between a choice being obligatory and such an ordering? This question seems to suggest that the moral status of an action, i.e., whether it is obligatory, is solely determined by it being better than or equal to other available choices. But this can't be quite correct; in fact, questioning this presupposition helps illuminate the second issue I have in mind.

Such a presupposition, i.e., that the moral status of actions can be solely determined by its *relative* 'goodness', treads upon a characteristic ambiguity in the informal concept of 'the moral nature of actions.' The issue here is how we can respect the distinction between choosing the better of two options, morally speaking, and a categorization of actions into good or bad, right and wrong, fair and unfair, or acceptable and unacceptable.[74] In what follows, I will be using the pair 'acceptable, unacceptable' as a generalization of all such categorization schemas. But as we will see, how one interprets the *semantic values* of the logic is a matter of choice, all else being equal.

---

[74] There may be other moral schemas, for example authentic or inauthentic, holy and debased c.f. [Williams 2012] pgs. 73-80.

While this might be controversial in some circles, I take the distinction between a ranking of choices and a categorization of actions to be part of the very bedrock of moral reasoning. Perhaps this is all too brief, but I think confusion regarding this distinction could be cleared up with the following considerations. Moral decision-making can often be quite simple; much of the time we can simply check to see how a certain action is correctly categorized. For example, in many cases, all we may need to know is that an action counts as *deception* and, as such, is (for example) wrong. However, in more complex situations, we may start with such a categorization of actions but cannot stop there—instead we must update how we think about the relevant actions given other options. To accurately represent this sort of reasoning, we need to answer a second question.

(C&MP)   How should a choice of one action over other options be related to the moral nature of the various actions?

There is still one more question that comes to the fore, especially with our focus on liberal theory. The issue here is closely related to O&O, the idea being how can we tell when some choice might be obligatory for one sort of person but not for another. In Chapter Three, I argued that the notion of a contrast class is insufficient to determine when an answer to an action-guiding question is satisfactory. We needed to also know what sort of answer the interlocutor would find relevant. The cases that I have in mind question what is fair to demand of such and such sort of person to do for the greater or common good, for example. The case of peace officers having much higher demands reasonably expected of them than, say, the general populace is especially germane. While some meta-ethics reject the idea of supererogation, we need a more general notion of

130

norms that we can make use of in liberal contexts; after all, liberalism is supposed to generally leave people alone, all else being equal. Given this approach to ordering options:

(SC)   How can we identify when it is unfair to oblige someone to choose the best option?

4.2 Choices and Options

In what follows, I argue that answers to these questions indicate that the notion of *obligation* should be formulated in terms of a complex relationship between *what option is best*, the *moral nature of actions* and the *limits to what we can reasonably expect of others*. In PRIT, this formulation is accomplished in part by combining an ordering of contrast classes with an ordering of actions with an ordering of what we can morally demand of different person-types or social roles. As such, this system uses a complex evaluation in its models. The evaluation I introduce orders three distinct primitive notions corresponding to the above answers: the moral acceptability of options, the relative moral superiority of decision-types and the moral limits to what we may demand of person-types. Ultimately, the hope is that with appropriate answers to these questions and a logic that enables us to reason with and about these moral orderings, we will be able to define a cogent deontic consequence relation. Given that the larger project is focused on using formal tools to investigate the nature of political principles and related concepts under conditions of pluralism, this will be a major hurdle to overcome.

I recognize that many philosophers will be familiar instead with evaluations that order one and only one primitive notion, i.e., truth (or some cognitive correlate), but

throughout this project I have defended, and will continue to do so, the idea that deontic reasoning is too complex to be represented using these more familiar accounts.

4.3 Best Choices and Unacceptable Options

To answer **O&O**, it is natural, if we set aside the question of what it means to ask too much of someone, to think of an obligatory choice in terms of which choice is best (morally speaking) given the circumstances. In other words, since obligations guide action, *each person should be guided by doing what is best given the circumstances* seems like a fitting notion with which to begin. My methodology will not only analyze the concept of the *best* but will also break down what we mean by "the circumstances" of a moral situation. There will be two parts to my formal treatment of relevant moral conditions or circumstances, namely specifying the options available and the type of person (or social role etc.) of the person making the choice. Thus, in order to analyze a notion of obligation apt for pluralist settings such as liberalism, I will develop an analysis of each component independent of each other and then combine the resulting structures to achieve my overall goal.

Before I deal with the notion of moral circumstances, it will be useful to distinguish three features of a moral ordering of choices that are relevant to developing a formal conception of obligation. We can extract criteria from these features that illuminate the surprisingly complex relationship between best choices and acceptable options.

The first feature I have in mind is that it looks possible to have many options in one particular contrast class be acceptable (i.e., to select) but from among these

acceptable options there may also be a best option. In other words, an action being *acceptable to choose* does not imply being the best choice unless we add other assumptions (in other words, the acceptability of an option does not guide action without further considerations). Again, we want our system to respect the distinction between a moral categorization of actions, in this case between acceptable to choose and not, and a comparative process for assigning moral status to *choices*.[75]

The second feature is that there may be ties for best in some moral orderings, and this is especially germane in the case of liberalism. Philosophically, this possibility indicates a subtle revision to what we can mean by *obligations guide action*: choosing what is best guides action when there are no ties for best. When we look for action-guidance, morally speaking, we want to learn what we are obliged to do and if, as suggested, we should be guided by doing what is best (given the relevant caveats) doing what is best (so construed) is what we are obliged to do and *vice versa*. However when there are such ties for best, a further choice, one that cannot be based on the moral reasoning that led to the tie for best, must be made. In other words, ties for best require a higher order notion of action-guiding than merely doing what is best or what we are obliged to do. After all, being told that there are two options we are obliged to do tells us nothing about which one to actually choose. Philosophically, this suggest that such ties

---

[75] In fact, as we will see, since in this system *weak permission* is defined in terms of *obligation* we can make a distinction between an *action* being *acceptable* to select and the *choice* of that action over alternatives being *permissible*. On this account, the notions of *obligation* and *permission* are formalized to account for the moral status of *choices*, while the notion of *acceptability* is used to partly account for the other (i.e., extensional) aspect of the moral nature of actions. While Scanlon has informed quite a bit of my approach, here we part ways. If my account is right, Scanlon's is problematic because he conflates a proposed action failing to be objectionable with its permissibility [Scanlon 1998].

for best could be used to represent moral dilemmas or disagreement regarding foundational policy in a liberal state; in fact, this is the approach I will take. On the formal side of things, these considerations suggest representing an *obligation* as a *maximal* element in a moral ordering rather than just as a *maximum* element, at least as a starting point for developing the concept of a morally best choice.

The third feature is that, for all that has been said so far, we could have an ordering that yields an obligation to choose an unacceptable option. This possibility seems, intuitively, quite problematic in a formalism for liberal foundations. The problem I have in mind is due to an ambiguity in how we use, correctly I assume, the moral imperative expressed by "you ought to…" locutions. Consider a moral ordering of choices, none of which is acceptable, but where at least one choice is better than the rest. If best choices are obligatory, we have an obligation to do something unacceptable. Perhaps even more worrisome: imagine if all available choices are not acceptable and an ordering posits that no choice is any better than any other option (or that a large subset ties for best of this bad lot). In this case, we seem to have a whole bunch of (trivially) best options, none of which is acceptable.

First, I want to head off a potential misunderstanding. I am not claiming that I don't need to represent that 'one should make the best out of a bad situation' in my system.[76] The issue here, as I see it, is whether such a choice, i.e., the best of a bad lot, should count as an *obligation* or rather something more along the lines of *the least you should do*. In what follows, I will develop and defend the latter idea; I call it the *obligation*/*approbation* distinction. Taking this distinction seriously will allow me to

---

[76] See definitions 4.16 and 4.21

134

distinguish the best choices we can use to represent *principles of justice* from the best choices we should sort in some other way, or so I will argue. I begin by clearing away some tempting lines of thought that may obscure this distinction.

4.4 Obligations and the Attribution of Responsibility

At a linguistic level, normally we have no problem with saying "since you *should*, morally speaking, do the lesser of two evils (all else being equal), you *ought* to do the lesser of two evils (all else being equal)." Now it may seem reasonable that if you *ought* to do the lesser of two evils, you have an *obligation* to do the lesser of two evils. But the hazy sort of obligation reflected in this linguistic phenomenon is not the right guide for our intuitions insofar as we are trying to formalize a notion of *obligation* that can provide the right framework for identifying *principles of justice*. Instead, that notion of *obligation* is closely tied up with the *responsibilities* we take ourselves to have.[77]

First, let me lay out a couple of preliminary definitions of the ideas at work here. The moral *should* of "you should do the lesser of two evils" I will call *approbatory* (i.e., you should do some evil that is *sanctioned* in this particular case). Secondly, recall that a choice that is *beyond the call of duty*, i.e., which asks too much of an individual to be required, is called *supererogatory*. In the context of an *obligation* we assume you are held responsible for the good or evil done (or foregone as the case may be), all else being equal. But in a *supererogatory* context, you are not held responsible for the good forgone but only the good you do, all else being equal. This basis for the

---

[77] Not to be confused with the similarly named abstraction norm from Chapter One.

supererogatory/obligatory distinction in turn illuminates a similar difference between the

obligatory and the merely approbatory.

In the case of doing the lesser of two evils, we tend to balk, and rightly so, at

ascribing responsibility to the unfortunate person for doing evil when she or he chooses

to do the lesser of two evils. For example, consider a parent caught in a typhoon with two

small children. If the parent holds on to both children, all three will die. However, one

child has a small chance of surviving if let go, while the other is still an infant and will

surely die if let go. Let us assume that the parent who lets go of the older child, in order

to have the best chance for all living, has done the best from available options, all else

being equal. However, it is also clear that letting go of either child is a bad thing, as is

holding on to both and damning all. But we generally, and correctly I assume, ascribe

responsibility to the parent in this case, not for the bad they do, i.e., letting go of the older

child, rather for the bad they avoid.[78] The lesson is that in an *approbatory* context, you

are held responsible for the evil mitigated, but not the evil perpetrated. This fact grounds

the obligatory/approbatory distinction in a way that clearly parallels the relation between

supererogatory and obligatory responsibilities; thus the cogency of the two distinctions

should rise or fall together.[79] And while in ethics and ordinary morality, some may

---

[78] As a caveat, such a parent is in a morally intolerable situation. Holding someone responsible for everyone dying because she or he didn't have the emotional wherewithal to actually let go of the older child seems cruel, or at least unduly harsh; doing so seems to blame a victim of horrible circumstances. However, ordinary responses to tragedy are not always the best guide to how we should formulate principles of responsibility, ethics etc.

[79] It might be objected that this analogy between *approbatory* and *supererogatory* is undermined by the asymmetry purportedly found between responsibilities for harm and good found in the Knobe effect (Knobe, J. 2003 "Intentional Action and Side Effects in Ordinary Language" *Analysis*, 63, 190-193). While an appropriate response deserves its own project, let me briefly sum up why I think the Knobe effect is

dispute the distinction between supererogatory and obligatory (some utilitarians or puritans perhaps?), in the context of political liberalism we need such a distinction— citizens can't be forced to give up, say religious beliefs, for a policy-makers' notion of the greater good.

With this analysis in mind, representing an obligation (*qua* best option) with the choosing of an unacceptable action undermines these distinctions. Moreover, it seems fitting then that the moral nature of actions, i.e., what sort of moral category actions belong to, should play a straightforward role in how *obligations* guide action. This suggests a simple change. We should instead understand *obligation* in terms of *best acceptable option*.

For current purposes, I will assume that an *obligation* is a best acceptable option, unless otherwise noted. Moreover, we will see that the same sort of formal structure that allows us to distinguish between what is supererogatory and what is obligatory can also be used to distinguish what is obligatory from what is merely approbatory (see definition 4.21). This suggests that the formalism accurately captures the distinctions between these different sort of best choices. In some sense, being able to formulate these differences in this manner should count as evidence that the *obligatory/approbatory* distinction is principled in general, and not just in the special case of the foundations of liberalism.

More to the present point, though, is that we have clear criteria for answering

O&O (How should we should identify which choices in a moral ordering are *obligatory*.)

---

illusory. The Knobe effect conflates the distinction between the deliberative and critical use of a moral principle (as elucidated by Scanlon 2008 see his pg 23 for definitions): conflating the *meaning* of a choice for an agent, i.e., the critical sense, with the *permissibility* of a choice, i.e., the deliberative sense, will lead to the same sort of mistakes Scanlon uncovers in the doctrine of the double effect. Thus if the doctrine of the double effect is illusory, so is the Knobe effect.

We therefore need our formulation of *obligation* to capture this relationship between *best choices* and the *moral status of actions*, i.e., **C&MP**.

4.5 Two Moral Orderings

With this preamble, it should not be surprising that our formal machinery takes as its starting point two distinct sorts of moral ordering—an ordering of *actions* and another of *contrast classes* or choice set-ups. We begin with some machinery for distinguishing *choices of actions* into acceptable and not acceptable.

Definition 4.6 (Course of Action) Let *A* be a set of action types. We will say that a finite sequence $\langle a_1, \ldots, a_n \rangle$, where each $a_j \in A$ is a *course of action* (or an *option*) from *A*. (Note that we assume $n \geq 1$, but *n* is finite, and we do *not* assume that $a_k \neq a_j$.)

Let *B* be the set of all courses of action from *A*, and **B** be the set of all finite subsets of *B*.

A *choice C* is an ordered pair consisting of a course of action *a* and a set of courses of action that includes it. That is, $C = \langle b, \varphi \rangle \in B \times \boldsymbol{B}$ such that $b \in \varphi$.

We can then designate the set of all choices over *A* by **C**. (We will use *a, b, c, d* etc. to designate courses of action, φ, ψ etc. to designate *sets* of courses of action.)

Definition 4.7 (Contrast Class) If $C = \langle b, \varphi \rangle$, φ is *contrast class* for *C*. That is, the tuples of action-tokens for the given choice is the *contrast class* for that choice.

Ultimately, for any *choice C* we want a structure representing whether that chosen element of the contrast class should be chosen (that is, whatever value the choice is assigned, it is as good as or better than any other option given). To define the valuation, we first need the points or nodes in our model.

Definition 4.8 (Situations) A *situation* is a set of choices. Let *Sit* be the [non-empty] set of all situations.

The intuitive notion here is that situations are the analogues for worlds in relational semantics for modal logic. We next introduce our first *valuation*, which will tell us which choices are *acceptable* and *not acceptable* (which we represent with a good old fashioned $\{1, 0\}$.)

Definition 4.9 (Valuation) A *valuation I* over choices in each situation is a function $I: \boldsymbol{C} \to 2^{Sit}$. This is equivalent to assigning a subset of *Sit* to each choice *c* by stipulating that for any $s \in Sit, s \in I(c)$, if $I(c)$ (*s*) = 1. We think of this as the set of situations in which *c* is acceptable. When *s* is in *I*(c), we also write $s \vDash^I c$ or $[\![c]\!]^I_s = 1$.

(We sometimes suppress mention of *I* or *s* where context makes it obvious which *I* or *s* is in question.)

Also note that a valuation assigns 1 or 0 to all choices for any situation; as it stands, there is no distinction yet between a choice being not acceptable because we have information that it is wrong, for example, versus not having enough reason to say that it is acceptable. In other words, there is no distinction yet between incompleteness and falsity; this comes later on when I introduce *negation* as a monadic operator, instead of as a semantic value.

Definition 4.10 We define a *pre-ordering*[80] $\leq_B$ on the options included in each contrast class $\varphi \in \boldsymbol{B}$, given an interpretation *I*, by setting for $\forall x, y \in \varphi, Cx\varphi \leq_B Cy\varphi$ *iff* in any situation where *I* deems $Cx\varphi$ acceptable, it also deems $Cy\varphi$ acceptable.

The idea here is as follows. Suppose $a, b, c, d \in \varphi$ in some choice situation *s*; both *a* and *b* may be acceptable options according to an interpretation *I*, that is $s \vDash Cb\varphi$ and $s \vDash Ca\varphi$. But if $Cb\varphi \leq_B Ca\varphi$ while $Ca\varphi \nleq_B Cb\varphi$ then (according to *I*) option *a* is morally better than is *b* even though both options are acceptable. We can now define frames and models for our choices in situations.

---

[80] A pre-order is a reflexive, transitive (but possibly non-total) binary relation.

Definition 4.11 (Frames and Models) A *frame* is an $F = \langle Sit, \leq_B \rangle$; A *model* is an

$M = \langle Sit, \leq_B, I \rangle$.

As previously argued, acceptability of an option is not enough to represent an *obligation*; we need to identify which options are maximal. We employ an obvious generalization of the previous notation to consider the effects of the interpretation and ordering on subsets of *Sit* and on particular choices. Given a structure $\langle \varphi, \leq_B, I \rangle$, we denote the set of maximal elements of φ according to $\leq_B$ with $Max(\varphi^I)$. If we let '$\boldsymbol{O}x\varphi$' represent that it is obligatory to choose x from φ we have the following first pass on defining *obligation*:

$$s \vDash \boldsymbol{O}x\varphi \leftrightarrow s \vDash Cx\varphi \ \& \ x \in Max(\varphi^I)$$

But as can be gleaned from the discussion of obligatory, approbatory and supererogatory distinctions, we still need a suitable answer to **C&MP**. And sorting out how the moral nature of the chosen action matters to the choice of it over other options requires us to identify the moral nature of each element of any contrast class. In some sense, this issue amounts to the question of what comes first, a *betterness* ordering similar to that above or the identification of the moral nature of actions.

Formally, identifying which option is best among acceptable options does not require identifying the moral nature of options from other contrast classes. And as argued earlier, the point of using a contrastive structure to regiment prescriptions (and cognates) is to represent how a choice's moral status is relative to a specifiable group of alternatives. It seems that with what we have so far, the moral nature of actions is merely derivative of

a betterness ordering; but this must be too fast, given the distinction between obligatory

and supererogatory and so on.

4.12 Better Choices

The order I defined above orders the options from one contrast class at a time. It

doesn't order choices drawn from different contrast classes, e.g., whether $Cx\varphi \leq Cy\delta$.

This distinction indicates that formally there is another moral ordering we need to take

into account.

My methodology uses a moral ranking of different *types* of decision-setups, i.e.,

contrast classes. The idea here is that we find ourselves frequently expressing a

preference for some sorts of decisions over others. For a non-moral example, choosing

which beer to drink from a selection of American beers is a clearly inferior situation to

one in which the selection includes beers from around the world. Similarly, on the moral

side of things, some types of choices are less morally ideal than others: Sophie's choice is

a clear example of a type of decision that is less desirable that almost any other type of

choice. Politically, a decision whether to bomb a hospital or a refugee camp (say in order

to stop a dirty bomb from being deployed) is again a type of choice no reasonable and

morally decent person would relish facing. In these cases, the options on the table as a

whole, i.e. the contrast class, are morally better or worse than a different class of options.

Definition 4.13 (Ideality-chains on contrast classes): Let *I* be an evaluation as above. An
ideality-chain for *I* is a tuple $\mathcal{B}^I = \langle \boldsymbol{B}, \preccurlyeq \rangle$ such that:
- $\boldsymbol{B}$ is a finite set of contrast classes as before
- We modify *I* to evaluate which contrast classes are morally better than or equal to
  others, i.e., $I: \boldsymbol{B} \to 2^{Sit}$. Here $\langle 1,0 \rangle$ represents, in the case of '1' an ideal decision
  setup and in the case of '0' a morally regrettable state of affairs, at least insofar as
  the options available when making a decision.

- $\preccurlyeq$ is a reflexive, transitive, partial binary relation on **B,** given by *I*.
- For all $\varphi, \delta \in \mathbf{B}, \delta \preccurlyeq \varphi$ *iff* $\forall s \in Sit,$ if $[\![\delta]\!]_s^I = 1$ then $[\![\varphi]\!]_s^I = 1$

where $[\![.\,]\!]^I$ again denotes the salient truth-set function.[81] We can then denote the set of maximal elements of an Ideality-chain according to *I* as $Max(\mathcal{B}^I)$.

We can now construct a composite ordering of choices together with this non-total pre-ordering of contrast classes. To create this new ordering we take the *coordinatewise ordering* on $\boldsymbol{C} \times \mathcal{B}^I$, a composite relation between the ordering of *choices* in $\boldsymbol{C}$ and our non-total, reflexive, and transitive ordering of contrast classes.

Definition 4.14 (Goodness-graph or G-graphs): *A goodness-graph for I is a tuple* $G^I = \langle G, \leq_G \rangle$ where $\leq_G$ is a partial pre-order on *G* such that*:*
- $G := \{g | g \in C \times \mathcal{B}^I\}$
- $(c_1, b_1) \leq_G (c_2, b_2) \leftrightarrow c_1 \leq_B c_2 \ and \ b_1 \preccurlyeq b_2$

So for each contrast class from $\mathcal{B}^I$, together with each choice within that contrast class, is now ordered by $\leq_G$. We again denote the set of maximal elements for G^I by $Max(G^I)$. A goodness ordering according to *I* seems a good approximation of a normative code; such a code specifies what is the best option[82] from each class of options, while each class of options is itself ordered. While perhaps not yet conclusive, with a *G-graph* we at least have a good handle on how to answer **C&MP**.

4.15 Obligations and Principles of Justice

At this point we have a fairly rough sketch of what our formulation of an obligation will look like. In other words, we can sort different best choices into different groups. But we still have no rule for identifying which sort of best choices should represent obligations.

---

[81] More precisely: the acceptable-set function.
[82] But is still only an approximation; see definition 4.16

For foundations of liberalism, we are interested in obligations so that we can reason about principles of justice, i.e. some class of obligations we will find interesting as principles of justice. Not only do we need to know what are the best options given a class of alternatives, but we need to know which sort of best options deserve to be seen as obligations and thereafter isolate some of these as our principles of justice. To get there we need to turn to issues of normative codes bumping up against an ugly reality.

We will need three operations on G^I in order to make sense of some common cases of moral-political reasoning, and thus to shed light on the nature of best options. The first helps capture the notion that we sometimes have to choose the lesser of two evils, or to make the best out of a bad situation (but not quite yet *approbatory*, see 4.21). The second is closely related and helps represent the notion that building castles in the sky is all well and good for theoretical purposes, but we should reject foolishly chasing idealistic options. The third allows us to represent how a code is transformed in the face of disagreement with other codes.

Definition 4.16 (Relative G-graphs) For any g-graph G^I, we want to represent what is best relative to a particular contrast class, say δ. We designate the greatest lower bound (with respect to $\leq_G$) of two contrast classes in G^I using $\wedge$, i.e. the meet. We define a *relative g-graph* $\boldsymbol{G^\delta} = \langle G^\delta, \leq_\delta \rangle$ such that*:*
- $G^\delta := \{\varphi_i \wedge \delta | \varphi_i \in G\}$
- $\leq_\delta := \{\langle \varphi_i \wedge \delta, \varphi_j \wedge \delta \rangle | (\varphi_i, \varphi_j) \in \leq_G\}$[83]

That is, the elements of the relative contrast class are just the meets of all the original elements with δ. An element of the new relative graph is less than or equal to another exactly when their non-relativized originals are so ordered. In other words, this technique lets us see "what happens, all else staying the same, if we assume that δ is the best

---

[83] This definition generalizes on a definition for a restricted P-sequence from [Benthem, Grossi, Liu] accessed Aug 7, 2013.

situation" by identifying δ with a top element, since it will be the top element of any contrast class ordered by $\leq_G$ and *ipso facto* also with $\leq_\delta$. We then define a maximal element for a relative g-graph in similar fashion from before. Thus, if we must choose what is best from among unpalatable options, for whatever reason, we can represent this with the relative g-graph for the contrast class that includes the unpalatable options. But as of yet, there is no need to identify such best options, i.e., from among unpalatable options, as obligations.

Not all sub-ideal reasoning involves choosing under bad conditions. In some cases, that which is more ideal than what one ought to do is instead impractical or infeasible. For example, a certain fox had a code that guided him to reach for some grapes; however, the grapes were out of reach and the fox replaced his old code with a new one. The new code reclassified the grapes as not so good after all—in fact, those grapes were probably sour. This sort of normative code change is more than just a change in which contrast class we take to be relevant. We need a more permanent type of change in a g-graph. In other words, we need an operation to chop off the unfeasible choices rather than just the "moving spotlight" operation of the relative g-graphs. But this cannot be a simple matter of lopping the tops off orderings since the changes that motivate removing upper options from an evaluation often ramify below as well: some options at a lower level of ideality will also disappear.

In the fox example, if we remove eating grapes as the thing most worth doing, the lower level option of wandering through vineyards will also likely be removed. Thus we cannot just take a relative maximal element as a new top element and keep the remaining

order the same to represent the sour-grapes phenomenon.[84] A more comprehensive

revision to I is required. And as may be gleaned from the fable, generally which options

are sour is a *material* rather than *formal* property of options. In what follows, it is more

perspicuous to use the strict part of our ideality ordering, i.e., $\prec$ instead of $\preccurlyeq$, in order to

make clear how sour-grapes reasoning works.

Definition 4.17 (Realistic Options and Sour-grapes): Let S designate a set of 'sour
options'. Given a G^I (a goodness-graph), we designate any $\delta \in \boldsymbol{B}$ a *realistic contrast
class iff* for any $\delta'' \in \boldsymbol{B}$, if $\delta'' >_I \delta$ then δ'' contains sour options, i.e., $\delta'' \cap S \neq \emptyset$. Thus
δ represents a 'second best' contrast class, i.e., a highest such class (according to I) that
does not include the "sour grapes" option(s). We will define G_{sg}^{δ}, the *sour-grapes* graph
for δ as follows. I° is a *sour grapes* reinterpretation of I *iff* I° is such that for any $\delta' \in$
$\boldsymbol{B}$ that is preferable to δ according to I, i.e., $\delta' >_I \delta$, $[\![\delta']\!]^{I^{\circ}}$ is non-comparable to $[\![\delta]\!]^{I^{\circ}}$.
G_{sg}^{δ} is G^{δ} defined relative to $G^{I^{\circ}}$ rather than G^I.

Note that it follows that $[\![\delta]\!]^{I^{\circ}} \in Max(\boldsymbol{B}^{I^{\circ}})$ since $\neg \exists \varphi \in \boldsymbol{B}$ such that $\varphi >_{I^{\circ}} \delta$; anything

previously thought better has been declared "sour grapes".

Since we order $\boldsymbol{B}$ in virtue of the membership of options and their ordering, a new

ordering of $\boldsymbol{B}$ will presuppose a revision of how some particular options are ordered.

Proposition 4.17.1 A sour-grapes interpretation, i.e., revision of $\boldsymbol{B}^I$, implies that some
choices that may have been comparable according to I are now incomparable according
to that sour-grapes interpretation.

---

[84] Please ignore any negative connotations associated with the use of 'sour grapes'.
While there are perhaps more frequent uses of this term in a negative sense, i.e., as a
sort of irrational distortion of preferences (maladaptive instead of merely adaptive),
psychologically it is generally helpful to 'pretend' that what you once wanted wasn't
really all that good anyways. For example, while Martha may be a fairly competent
martial artist and had many Sanseis encourage her to compete at world
championship tournaments, she discovered she is prone to injuries. Thus it seemed
to make sense to stop training for such a goal, and put her time and effort into her
academic work. And this seems eminently rational...and it is helpful if she convinces
herself that being a world champion would just mean she had to, say defend her title
until she lost it. And for Martha, the inevitable losing of her title would be worse for
her than winning it would be good.

Proof: Follows trivially from 4.14.

Thus, since membership is sufficient for discriminating contrast classes and we order contrast classes in virtue of the moral nature of the options, a change in the ordering of contrast classes presupposes a change in the ordering of individual options. The question is: how far can this revision go down the ordering? Intuitively, the idea is that if the reason why the option of walking in a vineyard would discriminate a better contrast class than one without that option is due to eating grapes being the most worthy thing to do (i.e., walking in the vineyard increases the prospects of finding grapes newly fallen to the ground, say, and so might rank ahead of more laborious options such as raiding the hen house). Once eating grapes is removed as the best possible outcome, it seems wandering in the vineyard could no longer be seen as better than other options, all else being equal. Since the detailed application of this idea requires appeal to the content of individual codes, i.e., that the value of some options is in terms of means for some other purpose, the notion above (4.17) can only pick out some formal constraints for a sour grapes interpretation. But as we will shortly see, this nonetheless allows us to demonstrate some fairly interesting results about the foundations of liberalism.

The sour-grapes definition illustrates that a moral evaluation is a complex system that depends both on an ordering of and specifying relevant choices set-ups as well as specific interpretations stating which choices are acceptable and which are not. Thus, which options are best is not a matter of logic (i.e., in this presentation, does not hold across all interpretations); and this is a virtue for a deontic logic modeling a plurality of moral codes. This leads us to our next topic, the difference between a moral code in general and a code that may be suitable in the foundations of liberalism.

146

4.18 Plurality of Codes and Liberalism

A plurality of codes results in at least some disagreement. As noted, a liberal society takes such a plurality as a starting point. In fact, calling it a 'society' involves the hope that the codes can (at least usually) get along despite the disagreement. With this in mind, we will need a way to represent what changes in a code when, sometimes at least, it can get along with others. Again, we will see some of the subtleties involved in using a *best option* to represent an obligation, especially in liberal foundations.

Definition 4.19 (liberal g-graph): Let L (for 'liberal') be the chain ⟨getting your way > compromise > recalcitrant disagreement⟩. A *liberal g-graph* for G^l is a coordinatewise ordering on $G^l{\times}L$, i.e., $⟨g, l⟩ \geq ⟨g', l'⟩$ *iff* $g \geq g'$ and $l \geq l'$.

As a motivating example, consider the simple (and overly simplistic) g-graph that says a society with no guns is better than a society with some guns. The relative g-graph for this code and L will rank the pair $⟨no\ guns, getting\ one's\ own\ way⟩$ best of all. Next best and not comparable to one another are $⟨getting\ your\ own\ way, some\ guns⟩$ and $⟨compromise, no\ guns⟩$. Next best but comparable with both is $⟨compromise, some\ guns⟩$ and so on…. In total we have a six-node directed graph corresponding to a coordinatewise ordering on this g-graph and the liberal L. A fact that will be important later on is that in virtue of this composite ordering with L there will always be at least one 'get your own way-pair' node better than any 'compromise-pair' node in a liberal g-graph.

Readers suspicious of formalism for its own sake might well wonder whether all this machinery is worth the trouble. Clearly, I think it is. Let us briefly turn to an example of its uses as a possible tool to reveal something about the foundations of liberalism, in particular how we can use the notion of a *best option* in this context. I will introduce two

more definitions. These definitions I take to encode important concepts central to

discussions of liberalism. If these definitions are apt then 4.19.3 proves an intriguing

claim about the moral nature of liberalism.

Definition 4.19.1 (liberal modus vivendi): A *modus vivendi* agreement is any
compromise-pair on a liberal g-graph that is less good than at least one getting your way-
pair.

Intuitively, a modus vivendi is an agreement for the time being. A modus vivendi implies

that at the earliest opportunity to get your way your code guides you to get your own way.

Definition 4.19.2 (liberal principled agreement): A *principled agreement* is any
compromise-pair of a liberal g-graph that is also a maximal element.

Proposition 4.19.3 A *sour grapes* interpretation that removes 'getting your way' as most
desirable is both sufficient and necessary for a principled liberal agreement.

Proof: Given a coordinatewise ordering in composing a liberal *g-graph*, it follows that
any liberal *g-graph* will have at least one 'get your own way'-pair as maximal. Thus
assuming no sour grapes operation is available, it follows any compromise-pair in a
liberal *g-graph* will satisfy *definition* 4.19.1, i.e. it will be a *modus vivendi*.
　　　　On the other hand, if we compose a *sour liberal g-graph* we can make a
compromise-pair of a liberal *g-graph* a maximal element and thus can satisfy *definition*
4.19.2. Here's how: assume we have the underlying liberal g-graph and its associated
interpretation *I*. Let $\delta \in L$ represent political compromise. We create a new graph G_{sg}^{δ} by
replacing *I* with I° such that we take the truth-set of any element of $L \succ_I \delta$ instead as
incomparable to $[\![\delta]\!]_{I^{\circ}}$ (with respect to set-theoretic inclusion). It follows that $[\![\delta]\!]_{I^{\circ}} \in$
$Max(\mathcal{B}^{I^{\circ}})$ *since* $\neg \exists \varphi \in \mathcal{B}$ *such that* $\varphi \succ_{I^{\circ}} \delta$. Thus, there is some $\delta^{I^{\circ}}$-pair that satisfies
definition 4.19.2.

The point of 4.19.3 is that what is best *simpliciter* is not obligatory; rather, in this case the

best compromise or consensus is obligatory. And so it should be in a pluralist, liberal

society. This *sour grapes*-style of change to a moral code is of fundamental importance to

the foundations of liberalism and the legitimacy of a liberal state, at least if we are

assuming a pluralist liberal society—more on this below. In section 4.42 I will show how

we can, with a minimal logic (i.e., one that has no rules besides the basic connectives and that by adding new rules could only prove more not less) prove a Rawls-ish and Scanlonesque criterion for allowing multiple, equally correct but conflicting moral codes (regarding policy). This criterion claims that being guided by an appropriate code implies that one must find grounds that those who hold disagreeing but nonetheless appropriate codes couldn't object to that choice. There I argue that the sour-grapes style of change to moral codes plays a fundamental role in interpreting this criterion.

4.20 More on the Different Types of Best Options

There is still more machinery required in order to make all the key distinctions we need. First, we will want to be able to distinguish what is *required* from what is *more than what is required.* Plausibly, one can think of circumstances in which, for instance, an action would be required of a police officer (or parent) but *not* a bystander. Thus we shall take as given a set *P* of 'types of person' or perhaps better 'social roles' and define obligation and related notions relative to such roles. Given the Cartesian product of the set of person types and maximal elements for some contrast class we can define a new predicate that can represent the distinction between obligatory and supererogatory, and, as promised, between obligatory and approbatory.

Definition 4.21 (Super-choices): We modify I to evaluate which maximal choices are too much to ask for each type of person, i.e., $I: P \times Max(G^\varphi) \to 2^{Sit}$. Here, $\langle 1,0 \rangle$ represents, in the case of '0', that the maximal choice is too much to ask with regards to that type of person, while '1' we interpret as coming under this limit. Thus, in our earlier discussion 'acceptable' was just shorthand that meant 'meets a minimum to be okay.' But once we no longer abstract away consideration of social roles, it becomes more natural to speak in terms of 'at least meets the limit.'

Given a set of person-types, P and the set of maximal choices from a relative g-graph, i.e., $Max(G^\varphi)$, we define a predicate Θ on $P \times Max(G^\varphi)$, such that for $c \in \Theta$, if $[\![c]\!]^I_s = 1$ and $[\![\Theta]\!]^I_s = 1$ we say that c is *obligatory* according to I for the relevant $p \in P$. And if $c \in \Theta$ and $[\![c]\!]^I_s = 1$ but $[\![\Theta]\!]^I_s = 0$ then *according to I, c is *supererogatory* for p; otherwise if $[\![c]\!]^I_s = 0$ but $[\![\Theta]\!]^I_s = 1$ for $c \in \Theta$, then c is the best of unacceptable choices (the least you could do: *approbatory* according to I).

Intuitively the super-choice structure tells us which maximal choices are unacceptable

relative to prescribing said choice for a particular person-type. And as can be seen, it also

represents which choices, normally objectionable, become sanctioned when no other

choice is better. Alternatively, we can interpret Θ as telling us whether the choice to

prescribe a particular maximal choice for some person is acceptable or not.[85] On this

account, obligations amount to which maximally acceptable choices may be (i.e.,

unobjectionably) prescribed for a certain type of person according to I.

There is another, more political, distinction that we will want to represent. In

4.19.2 I defined the notion of a principled agreement. But Martha may find herself

making a principled agreement, i.e., one that she is committed to regardless of better (in

terms of, for example, self-interest) options coming by, even though there is no moral

principle guiding her to do so. She may be a psychopath who understands the long-term

but self-interested benefits of such principled compromises and cooperation (no political

affiliation should be read off this description). On the other hand, Mary's moral code

regarding political behavior guides her to always find compromises that she commits to,

all things being equal. And lastly, John's moral (so-called) code tells him to make

political compromises in the face of recalcitrant disagreement, but to try to get his own

way at the earliest opportunity. Thus we could be principled in compromising in the first

---

[85] Or both, as possible in the formalism of *RD-choice*; if this is a desirable possibility, set any maximal choice that is non-objectionable relative to a person-type as obligatory instead.

place, but that the compromise fails to be itself principled, i.e., it's merely a *modus vivendi*. The philosophical pay-off for these distinctions is that, of course, as moral agents in the political sphere our ambition is to reach principled compromises for which the parties to the agreement were principled in doing so. And the idea here, once again, is that a formal representation may serve a fruitful role in analysis.

In order to represent the distinction between a principled compromise and being principled in compromising in the first place, we can use the notion of a *strong permission*. The notion of a *weak permission* is just that there is no obligation to refrain from the permitted choice. A *strong permission* can be understood as permitting an exception to a norm, i.e., while there is an obligation to refrain from that choice, the force of this obligation is set aside for a different obligation (for whatever reason). A strong permission tells you that under particular conditions, one need not be guided by the relevant norm. Being able to represent the distinction between a weak and a strong sense of *permission* may, in general, be a useful feature in a deontic system. It just so happens that I will represent the notion of being principled in compromising in the first place as an instance of a strong permission.

It may be helpful to understand this weak/strong permission distinction by appealing to a particularly literal notion of what it means for some claim to be morally action-guiding. On this account, only prescriptions can be morally action-guiding; they prescribe what one should do and, as in the case of injunctions and directives, are intended to change peoples' behavior. The idea here is that one can guide action only if it is possible to change someone's behavior by said guidance. A weak permission has no such (perlocutionary) force, all else being equal: if someone apparently changes her

behavior due to a weak permission, this change presupposes independent motivation, morally speaking, for the behavior. In other words, the agent had reasons to do *X* and thought that she shouldn't choose to do *X* but then discovered that, after all, she was permitted to do *X*. On this account of action-guiding, it is the agent's prior motivation for *X*ing that accounts for her choice, not the following of a moral code. The strong permission is thus morally action-guiding in a way that the weak fails to be. The strong permission is intended to change someone's behavior by changing what norm they should follow. While perhaps not all may find it necessary to tie in action-guiding so closely with the living up to norms, the distinction between weak and strong permission is not without further philosophical pay-off. As we will see, representing the notion of trumping as a change in contrast class leads naturally to formalizing this notion in terms of a strong permission—a change in which norms apply is a change in which maximals we appeal to, and which maximals we appeal to depends on the relevant contrast class.

In PRIT, I will represent a strong permission as guiding action by specifying a change in contrast class. In particular, a strong permission tells you when you can ignore best options and then choose best from the remaining acceptable options, i.e., the second best from the original class. Thus, as in other cases of sub-ideal reasoning, by changing the contrast class, we are then guided by doing what is best from that class of options. To help clarify the connection between strong permission and action-guidance, we can use some of the surplus structure (that is, surplus in respect to what is obligatory) found in a relative g-graph.

Definition 4.22 (Strong permission): We interpret every $[\![Cx\varphi]\!]_s = 1, s.t. Cx\varphi \notin Max(G^\varphi)$ as an element of a *strong permission class*. To distinguish obligatory maximal choices from supererogatory maximal choices, we interpret every $[\![Cx\varphi]\!]_s = 1$ such that

Cxφ ∉ Θ as an element of a *strong permission class.* We then use the relative g-graph operation to identify what is best from this strong permission class. A *strong permission* is a norm calling for us to choose a maximal from this strong permission class.

A strong permission may be used to represent a norm calling for compromise when sticking to your guns will result in recalcitrant disagreement, but nonetheless allowing for the compromise to be a *modus vivendi*. This is again a case of demonstrating the limits to the formal properties of such agreements and where material considerations must enter. As we wanted, a strong permission is action-guiding in the sense explained above: it tells you to choose a maximal from a contrast class and as such, is intended to possibly change an agent's behavior, all else being equal.

In fact, on this account one needs a strong permission that says whether it is permissible to trade in your code for a sour version; that is, if you ought to compromise, a strong permission could tell you it's okay to give up on your best options. Thus we can make a distinction between agreements that are principled (i.e., 4.19.2) and being principled in compromising in the first place (i.e., 4.22).

The more familiar notion of permission, i.e., what we call a *weak* permission defined in terms of obligation, will be represented by *not being obligated to not choose*. The formal representation for this sense of permission must wait until the formal language and its operators are introduced.

4.23 Representing Liberal Choices and Codes

We can, with the tools developed so far, make a first pass on representing some basic notions in the foundations of liberalism. To illustrate, we can represent a minimal notion of liberty in terms of that which is too much to demand of a privileged person-

type: the citizen. Intuitively, we want to distinguish between maximal choices a liberal

state can demand a citizen to do for, say, the ruling party's idea of the greater good and

maximal choices that are, generally, unfair to demand of ordinary citizens.

Philosophically, the idea here is that liberal states aim to leave people alone to live their

lives as they see best, all else being equal. The reason why I call this a minimal sense of

liberty is that the conditions regarding who counts as a citizen are essential to recognizing

a state as appropriately liberal, i.e., states with second class citizens (e.g., those who don't

have the same standing in the law as first class citizens) generally don't deserve honest

title to *liberal*. While there is more to be said about the concept of liberty, for now I can

develop this minimal notion as a sub-sort of super-choice.

Definition 4.24 (simple liberty): We use $M \in P$ to denote the type "citizen." We use the extended evaluation from 4.21 (Super-choice) and define a predicate Ξ on $M \times Max(G^\varphi)$ such that for any $c \in \Xi$ if $[\![c]\!]_s^I = 1$ and $[\![\Xi]\!]_s^I = 1$ we define c as *obligatory for all citizens* according to *I*. If $c \in \Xi$ and $[\![c]\!]_s^I = 1$ but $[\![\Xi]\!]_s^I = 0$, according to *I*, *c* is *illiberal*.

Obviously, there is conceptual space in the simple liberty structure for a correlate to

approbatory as in super choices; I will leave this notion informal, with the unimaginative

name of 'liberal approbatory.' One of the goals for PRIT is to sort out what it means to

cooperate while leaving others alone to disagree with us; the formal condition above, i.e.,

for illiberal policy, won't get us very far without assuming this material property of

liberal societies. The natural thing to do here is to combine the techniques from

representing compromising in the face of disagreement with the above structure for

representing leaving people alone.

First I define an ordering of simple liberty choices, constructing a g-graph where all options are considered in terms of what is fair to ask of citizens: instead of using $\leq_B$(4.14) we will define an ordering with respect to Ξ. We then compose this ordering with L as we did with liberal g-graphs (4.19). Finally, to complete the code, we turn it into a sour ordering.

Definition 4.24.1 (Ξ g-graphs): A Ξ *g-graph* is a $\Xi^{\preccurlyeq}$ such that for each $\langle c, \delta \rangle \in \Xi$, $(c_1, \delta_1) \preccurlyeq_\Xi (c_2, \delta_2)$ *iff* according to I from 4.24 for any s, if (c_1, δ_1) meets our liberal minimum at s then so does (c_2, δ_2).

Definition 4.24.2 (liberal $\Xi^{\preccurlyeq}$): A *fair-liberal graph* is the coordinatewise ordering of $\Xi^{\preccurlyeq} \times L$.

Next, I treat the call for political compromise in a political sphere chock full of disagreeing codes as an instance of a sour grapes style of change to fair-liberal graphs.

Definition 4.24.3 (sour liberal $\Xi^{\preccurlyeq}$): Let S designate a set of 'get your own way' options (our sour grapes, so to speak). Given a fair-liberal graph $\Xi^{\preccurlyeq}$, we designate any $\delta \in B$ a *realistic contrast class iff* for any $\delta'' \in B$, if $\delta'' \succ_I \delta$ then δ'' contains 'get your own way' options, i.e., $\delta'' \cap S \neq \emptyset$. We will define Ξ^δ_{sg}, the *sour-grapes fair-liberal* graph for δ as follows. I° is a *sour grapes* reinterpretation of I *iff* I° is such that for any $\delta' \in B$ that is preferable to δ according to I, i.e., $\delta' \succ_I \delta$, $[\![\delta']\!]^{I^\circ}$ is non-comparable to $[\![\delta]\!]^{I^\circ}$.

Thus, we interpret δ to represent the class of choices that result in political compromise.

With 4.24-4.24.3 we can represent a liberal code that guides the choice of policy in terms of prescribing the best upon which we can agree, subject to protecting the liberty of individual citizens. In order to represent being principled in compromising, we can straightforwardly compose a strong permission in the same way as before.

With this logical machinery, we can represent sour-fair liberal codes that could disagree with each other regarding what we mean by the 'best', others upon the meaning of 'liberty' and still others will disagree on how we should reach a compromise. Thus we

must distinguish between what a code says is a best acceptable choice, i.e., obligatory according to that code, and what is a best option acceptable to all relevant codes. To handle all these codes and the resulting disagreement, we need a formal language and its proof theory to reason about all these different orderings.

4.25 The Proof Theory—Preliminaries

Often it would be sufficient to have a semantics in hand, especially when the primary goal is to apply logical insights to issues in other areas of philosophy. However, in our case, it is also important that we develop a useful proof theory. The idea here is that with PRIT foundations, we care not only about *what* we can prove (which we might be able to get at merely by considering the class of valid formulas), but also about *how* we might prove our claims. And just providing a procedure for deciding whether a formula is valid won't be good enough in our case; we need tools to analyze proofs. Analyzing proofs themselves will allow us to identify 'parameters' used in a sequent, i.e., a mathematical object expressing a conditional judgment, and investigate what sort of changes we could make to these parameters if want to see different conclusions. This sort of approach allows us to focus on problem solving; we can see what went into proving a conclusion we need to avoid (like liberalism being incoherent if there are multiple correct accounts of what is just) and sort how to make suitable changes (like making a distinction between inconsistencies that are disastrous from those that can be reasoned with fruitfully in a pluralist setting like liberalism).

The other reason to have this sort of proof-theory for PRIT is that sometimes we might want to keep track of, i.e., recover information glossed over at a more general level of abstraction, which moral theory was used to warrant some course of action. And to be

156

slightly more pragmatic, I have also tried to get to interesting proofs, i.e., with meta-ethical import, as directly as possible. You will find in this section a proof claiming that, in a pluralist setting, moral codes must refer to and require compatibility with the content of competing codes. Not only does this proof bear a resemblance to Rawls' claims about the implications of what he calls 'the fact of Reasonable Pluralism', I prove this version of the claim using none of his theoretical assumptions or framework. In fact, this claim is proved merely in virtue of a weakest logic capable of defining the relevant logical words. This proof can be found at 4.42.

In keeping with two of the recurrent themes, I propose we want a proof-theory for our deontic *action-guiding concepts* in terms of a formal language that sets out how some *sorts* of choices logically follow from other choices. But this phrasing is slightly misleading. Granted, on a direct and simple reading, it is the choice of actions for which we seek moral guidance; the issue here however is that we need more than just a typing of *choices* in order to represent reasoning about, for example, *conjunctions* of choices. That is, a conjunction of choices is not itself automatically a choice. For example, the conjunction of a choice to lower spending on foreign aid (instead of, say, keeping it steady or raising it) and the choice to lower taxes for those who tend to vote for your party (instead of, say, abolishing regressive tax breaks, i.e., tax breaks that required a high income in the first place) is not itself a choice: choices are structures with contrast classes and there is no contrast class with this conjunction, as it stands. Moreover, it's not clear how a logic could automatically populate a contrast class for such a conjunctions,

i.e., extensional conjunctions, in a coherent manner.[86] Still less are disjunctions of choices

themselves choices, let alone arbitrary conditionals. Since we want to know how certain

types of things follow from other types, the concept of *choice* is not apt, on its own, for a

type-theoretic approach. At the very least, we should look for a sort of typing that is

simpler to work with.

The lesson here is similar to a point raised in Chapter One. There is such a thing

as too much detail for some purposes [Sambin & Valentini 1998]. Recall how in order to

do some maths, some details important at a fundamental level must be ignored or

abstracted away, i.e., too much information can be "a burden when dealing with the

synthetic methods of mathematics (pg. 2)." And again, what makes such abstraction

constructive is that such details can be recovered when warranted. In like fashion, the

fine-grained structure of *choices* is important to solve problems at a fundamental level.

For example, when the order or priority of actions matters, I have urged the use of

something like the fine-grained structure of *choices* to represent the relevant reasoning.

However, when such issues are not in play, and we are concerned (for example) with

conditional obligations or the combination of choices, the added structure of *choices* is

cumbersome (to say the least). In such cases, I propose we can treat the topic action of a

choice *as if* it were a 'formula' on its own right. Alternatively, one can think of this

---

[86] The closest I have been able to fix this issue is that each possible pair of options, one from each component choice, populates the contrast class for a conjunction of two choices. The selected option is thus also a pair. But this sort of conjunction is not very extensional in nature (we must reference other options not selected) and thus not apt for an extensional notion of conjunction, i.e., it is incoherent if a choice is an extensional conjunction. Later I will introduce an intensional notion of conjunction called 'fusion' and this seems more like what could be going on with this sort of conjoining of two choices.

seemingly exclusive emphasis on the topic as actually elliptical for our discussion of the corresponding choices.

I propose then that we should interpret our formulas as plans. One of the characteristic features of moral and political reasoning, i.e., in terms of justifying policy, is the role of commitments. It is our commitments to make certain choices under certain conditions— where we foresee selfish or self-interested confounding factors— which is, in some sense, paradigmatically the subject matter of moral reasoning and related action-guiding concepts[87] (i.e., instead of say, prudential reasoning where cognitive updating of new information is paramount). Such commitments can be described in terms of [moral] *plans*. I propose we do so; consider some of the following examples of such typing.

4.26 The Types—Plans

The question in front of us is how to formalize logical complexes of choices and at this point, how to interpret what such complexes even are, since as noted they don't seem to be choices in any straightforward sense. My suggestion has been to treat them, instead, as *plans*. Let us begin, then, with suitable interpretations of atomic formulas. A choice is, in a natural if not a rather simplistic sense, a plan. My plan for the next few minutes may be that I will choose a green tea instead of another cup of coffee. And if we focus on the topic of choices, this is a plan to drink green tea. The advantage of talk of plans rather than choices is that such talk more readily generalizes when we bring plans together via logical operations.

---

[87] There is here, I think, a very interesting philosophical moment. On this picture, moral reasoning implies the refusal to countenance updating the contrast class of a choice under *ordinary* conditions. Now what counts as *ordinary* conditions is what is so philosophically intriguing. But such issues are beyond the scope of this project.

Suppose we have a *plan* to choose *A* and the same plan is also a plan to choose *B*, it is thereby a *plan* to choose *A* and *B*. A plan to buy a car that is also a plan to give away all my money involves, if it is a smart plan, ordering my choices in such a way that both will be achieved. It must be a plan to buy a car *and* give away my money.[88] The case for disjunctions is similar: while in standard propositional languages *A* implies *A* or *B*, the disjunction of a choice to *A* or a choice to *B* is not, necessarily, itself a choice. But a *plan* to choose *A* is also a *plan* to choose *A* or *B*, extensionally speaking.[89] Similarly, a plan is *type A or B* just in case it is either a plan of type *A* or a plan of type *B*.

Negation is an interesting connective, for both plans and choices. Certainly the negation of a choice is not naturally regarded as a choice—not choosing armed rebellion and choosing not to rebel are two quite different things, the former occurring much more frequently than the latter. Similar remarks apply to plans. The distinction between choices and plans regarding negation comes, instead, because we must allow for negation of complex formulas and not just atoms, and complex formulas, as such, can be plans but not choices.

---

[88] And as argued in the last chapter, questions of ordering actions is a question of how to choose which action to do in which order. So we distinguish between smart plans that can achieve both goals and ill conceived plans in terms of choices that order the options differently.

[89] However, it is also possible that we want a more nuanced (i.e., multiplicative or intensional) notion of disjunction. For example, suppose that today we don't know for certain which options will be in play tomorrow. Let's assume that there are two groups of options, φ and σ we are concerned with; the plan is that if φ is in play choose *A* from it and if σ is relevant, choose *B* from that group. Thus we have a plan to choose *A* or a plan to choose *B*. That is, to represent a disjunction of *plans* may require keeping track of which plan matters to which choice. While this may be a useful feature to have in a deontic logic, i.e., if we want to go a non-distributive route, I will not pursue it any further, as it would meet no pressing need to do so in the immediate project.

Since not planning is not itself a plan, all things being equal, we will interpret the negation of a plan *A* to be a *plan not to A,* i.e., as a plan to do something incompatible with doing *A*. One naturally thinks of "not *A*" as the disjunction of all the plans that are incompatible with doing *A,* so for instance planning not to engage in armed rebellion is the disjunction of planning to lead a quiet life or planning for a career in government and so on.

It will be useful to our overall project to have a negation built on the notion of incompatibility. We want to allow for reasoning about cases that can't be actualized, notably ideal cases and about plans that can't be jointly acted upon. A straightforward and germane case is when two different groups in a political community have different plans for a certain chunk of tax revenue and where if one group gets their way the other will lose out. Understanding deontic negation in terms of *incompatibility* rather than in the more simple terms of "*not*" allows us to readily distinguish that case from the situation in which one group has plans for the tax revenue and another hasn't ever thought about it and really doesn't care what happens.

Before I set out conditions for implication, we need to be able to distinguish cases where the same plan guides us to choose both *A* and *B*, and where we combine a plan to choose *A* with a different plan to choose *B*. In relevant logics, this second sort of conjunction, i.e., *intensional conjunction*, is called *fusion*. We say a plan is of type *A fuse B* if it is a plan to combine a plan to choose *A* with a plan to choose *B*. This is in contrast with the extensional conjunction case; as explained above, there we say a plan is type *A and B* if it is a plan to choose *A* and *B*.

In the proof-theory for relevant and associated substructural logics, implication is generally defined in terms of *fusion*. Let me give a brief example of how this relationship works. Suppose we have a plan, a type *A fuse B* plan, and our theory tells us that it is therefore also a type *C* plan. This means that if we have an *A* plan, it follows it is also a plan to choose a type *C* whenever *B* is chosen, i.e., that 'choosing *B*' implies 'choosing *C*' follows from choosing *A*. However, the connection between antecedent and consequent in relevant implication is subtler and more sophisticated than may perhaps be thought from this brief characterization. And this connection is due to the constraints or properties we assign to *fusion*.

4.27 Proof Constituents: Sequents, Structures and Punctuation

In an effort to describe the system in an efficient and readable way, I will leave aside certain details I assume most philosophical readers will be familiar with. The system may be unfamiliar in that it includes 'rules' of two different sorts. First, each logical operator is equipped with introduction and elimination rules of a sort I assume is familiar. The feature that may be less familiar is the inclusion of *structural* rules. Structural rules are, intuitively, ways to pay closer attention to the premises (or evidence) by regarding them as a sort of structured entity, whereas in the more common systems, the premises are treated as an unstructured collection. For example, in proof systems designed for classical or intuitionistic logic, we generally find that the premises are represented as belonging to a set, whereas in the proof theory we will be looking at how the order of premises could matter or that a premise couldn't be used more times then it

162

appears, for example. So we will need to add structure to sets, so to speak,[90] to represent

how we could use premises in these different ways. In order to make this idea clear, let

me introduce the notions of a *structure* and *punctuation*. (In what follows I am assuming

the notion of a *language* in terms of a set *Lang* composed of its atomic formulas and

connectives that satisfies unique decomposition. [Restall 2000], see his pgs. 14-16.) I will

use letters like "*A*", "*B*" and *"C"* for formulas and reserve letters such as *"X*", "*Y*" and "Z*"*

for structures, except of course where formulas are being used as structures.[91]

Definition 4.28 (Punctuation marks)
An object together with its arity *n* (the number of input arguments) is a *punctuation mark*
p when its arguments are structures.

Definition 4.29 (Structures)
Given a set of punctuation marks *Punct,* disjoint from *Lang*, the multiset *Struct*(*Lang*,
Punct) will be a collection of *structures* when it satisfies the following unique
decomposition condition.

Lemma 4.29.1(Unique decomposition for structures)
$Suppose\ X \in Struct(Lang, Punct)\ and\ X = pX_1 \dots X_n = qY_1 \dots Y_m, thus\ p = q, n$
$= m\ and\ for\ each\ i = 1, \dots, n, X_i = Y_i$

Like *Lang*, if *Struct* satisfies unique decomposition, we can use induction to prove claims

about structures. We can think of structures and their punctuation as a way of

representing how we combine, compose and manipulate premises in a conditional proof.

In order to show unique decomposition, we need to make explicit some further

constraints on *Struct*.

---

[90] The conceptual priority here is assumed merely for expository purposes. We
might instead think of sets where the structure of, say, multisets, has been stripped
away.
[91] What follows is adapted from Restall [2000], though the proofs regarding the
unique decomposition of structures and induction over structures are my own.

Lemma 4.29.2 (Struct conditions)
Let Struct be the smallest subset of $(Lang \cup Punct)$ *such that* $Lang \subseteq Struct$ *and if* $n(p) \in Punct, then \forall X_1 \ldots X_n \in Struct, pX_1 \ldots X_n \in Struct$

Lemma 4.29.2 states that since any formula is a structure, all other structures are

constructed by combining these structures, i.e., the formula, using punctuation, similar to

the case of atomic formula and connectives in *Lang*.

Proof for Lemma 4.29.1*:* Now follows straightforwardly from the unique decomposition
of formula in *Lang*.

Theorem 4.29.3 (Induction over structures)
Take any collection of structures Struct(*Lang, Punct*). *Suppose the property* Ψ *holds of
every formula of Lang* (our base case for *structures*) *and suppose that for each
punctuation mark* $p \in Punct$ *with arity n, whenever* Ψ *holds of each* $X_1 \ldots X_n$, *then* Ψ
holds for $pX_1 \ldots X_n$ *as well. It follows that* Ψ *holds for each structure in Struct*(*Lang,
Punct*).

Proof: We assume the collection Q of all formula satisfying Ψ. Q must also satisfy the
conditions in Lemma 1.3 since $Lang \subseteq Struct$. But that means that *Struct*(*Lang,
Punct*)$\subseteq Q$, since *Struct* is the smallest such collection. Therefore, every structure in
Struct(*Lang, Punct*) satisfies Ψ as well.

We will also use the fact that, just as formulas have subformulas, structures have

substructures.

Definition 4.29.4 (Substructures)
We define the *substructures* of a structure as denoting the elements of the smallest set of
structures satisfying two conditions:
   • Any structure X is a substructure of itself
   • If X is $p(X_1, \ldots, X_{a(p)})$ then every substructure of any of the structures X_i is also a
     substructure of X.
I will write '$Y(X)$' to indicate a structure, Y, with X as its substructure.

Definition 4.30 (Antecedents, Consequents and Consecutions)
Given *Struct*(*Lang; Punct*), we define $\vdash$ to be the *consecution* relation on $Struct \times Lang$
such that $X \vdash A$, where the *antecedent* is an $X \in Struct$ and the *consequent* is an
$A \in Lang$.

Before we can define when we have a proof of a consecution, we need the notions of

inference, *rules*, and *proofs*.

Definition 4.31 (Inferences and Rules)
The pair *Inf* = $\langle(\vdash),\vdash\rangle$ is an *inference* such that $(\vdash)$ is the set of consecutions that are premises and the latter of the pair, i.e., the single consecution, is the conclusion. We define the set of any such pairs *Rule* = $(\langle(\vdash),\vdash\rangle)$ to be the *rule* corresponding to the inferences.

There is a particular sort of rule that is central to type-theoretic proof theory. Intuitively,

structural rules let you rearrange or adjust structures while not 'losing' any of the

formulas that followed from the antecedent structure. Before we can formally introduce

the notion of *structural rules*, I need to explain what I mean by *closed under substitution*

for formulas. As you can see from above, structures (uniquely) decompose into formulas;

now, if a rule requires that you can replace any formula in a structure with an arbitrary

structure, that rule is closed under substitution for formulas. Also, when a structure, say Z,

is generated *via* a structural rule from X, we will denote this relation by $X \Leftarrow X'$, where in

this example X' stands for Z (given that the consequent with X as antecedent and the

consequent with X' is the same).

Definition 4.32 (Structural rules)
A rule $(\langle(\vdash),\vdash\rangle)$ closed under substitution for formulas is a *structural rule* when the structure of the conclusion is an X' and the structure of the antecedent is an X such that $X \Leftarrow X'$.

Definition 4.33 (Rule families)
We define a collection $\{X|(\langle(\vdash),\vdash\rangle) \in Struct(Lang;Punct)\}$ to be the *Rules* for a *Lang*.

Typically we will want one rule for introducing and one rule for eliminating each

connective in that language. Also, for each punctuation mark and given any structure

composed with that punctuation, there is a formula with a connective equivalent to that

structure in virtue of 4.29.2 (i.e., in virtue of the unique decomposition of structures.) Finally, any structural rule governing each punctuation mark is a member of *Rules*.

4.34 Proof theory with no structural rules

Adding structural rules lets us prove more, not less. So, before I delve into the intricacies of the particular proof theory and formal semantics, let me first prove an interesting claim that must be true of any more complicated system. This claim is interesting because, while it is a valid deduction in a weakest such proof system (i.e., given suitable elements for representing the relevant deontic notions), it makes what may first strike one as an ampliative claim when interpreted back into ordinary language: it makes an inference from what our logical words mean to what we will find in moral codes acceptable in any liberal society. Obviously, this deduction is not ampliative. Instead, it seems best to interpret it as revealing a requirement of any moral code under conditions of pluralism: that any such code must guarantee that acting upon its prescriptions is compatible with the obligations of rival codes. This criterion bears a strong resemblance to a claim Rawls infers from what he calls "the fact of *Reasonable pluralism*." What this proof shows is that this conclusion holds without using any of his theoretical apparatus or assumptions.

Let me now introduce the rules for our connectives. As noted above, these are the mostly familiar introduction and elimination sort for each connective, though as I have defined rules, we could have a rule with no premises. And this is how I will introduce the *identity axiom*: as a rule with no premises.

Definition 4.35 (Rule for Identity)
The *rule for identity* is the collection of all inferences with no premises to a conclusion of the form $A \vdash A$

The connectives we will now be looking at will be: implication, fusion, conjunction, disjunction, negation and obligation. There is a small technical caveat here. I promised to start with a deontic logic with no structural rules, and if I were to be accurate in how I present such a logic, I would have to define two different implications, the familiar left-to-right arrow and a right-to-left arrow. It is only in the presence of a structural rule letting us ignore the order in which premises may be combined where these two sorts of arrow need no longer be distinguished. We find this sort of structural rule in the familiar classical and intuitionistic logics, as well as the standard take on relevant logic, hence it is not surprising that the left-to-right arrow is relatively well-known compared to its backwards sibling. For ease of presentation and also due to the lack of any present need, we will ignore the right-to-left conditional in what follows. And this same line of reasoning applies to the fact that we would also need two different sorts of negation in any logic without structural rules.

We begin with the rules for the conditional. They are the familiar rules of modus ponens and conditional proof, though written in sequent notation. Note that in the case of conditional elimination, in the conclusion itself we explicitly keep track of all the evidence required to arrive at the conclusion B by including both X and Y in the antecedent of the sequent.

Definition 4.36 (Implication Introduction and Elimination Rules)

$\rightarrow I$

$$\frac{X; A \vdash B}{X \vdash A \rightarrow B}$$

$\rightarrow E$

$$\frac{X \vdash A \rightarrow B \quad Y \vdash A}{X;Y \vdash B}$$

I briefly mentioned above that the key difference between a relevant and a classical or intuitionistic conditional is that the arrow has precisely the relationship these rules stipulate for the semi-colon to different operators, viz. to fusion for the relevant arrow and to conjunction for the classical or intuitionistic arrow. I introduce a bit of technical vocabulary for this relationship.

Definition 4.36.1 (Punctuation residuating Conditionals)
We say a connective $\rightarrow$ is a *conditional* that residuates the punctuation mark ';' *iff* it satisfies $\rightarrow I$ and $\rightarrow E$.

I now stipulate the rules for the fusion operator that ensures that it is related to semi-colon in the appropriate way.

Definition 4.37 (Fusion Introduction and Elimination Rules)

$\circ I$

$$\frac{X \vdash A \quad Y \vdash B}{X;Y \vdash A \circ B}$$

$\circ E$

$$\frac{X \vdash A \circ B \quad Y(A;B) \vdash C}{Y(X) \vdash C}$$

I have suggested above that fusion is "another sort of 'and'." It's not obvious what this new sense amounts to, glossed in ordinary language terms. For the present, it might be intuitive enough to see fusion as 'together with' and in some cases, 'compatible with', 'composed with' or 'combined with'. These interpretations of fusion will be appealed to

168

in 4.42. Let me now introduce the more familiar formal notion of 'and,' namely conjunction.

Definition 4.38 (Conjunction Introduction and Elimination rules)

$\wedge I$

$$\frac{X \vdash A \quad X \vdash B}{X \vdash A \wedge B}$$

$\wedge E_L$

$$\frac{X \vdash A \wedge B}{X \vdash A}$$

$\wedge E_R$

$$\frac{X \vdash A \wedge B}{X \vdash B}$$

The rules for disjunction for disjunction are familiar; thought the form that disjunction elimination takes may require some unpacking. It may help to think of the two leftmost premises in relation as establishing that each of A and B, together with whatever, if any, subsidiary information is provided by the rest of Y, are enough to allow us to establish C.

Definition 4.39 (Disjunction Introduction and Elimination Rules)

$\vee I_1$

$$\frac{X \vdash A}{X \vdash A \vee B}$$

$\vee I_2$

$$\frac{X \vdash B}{X \vdash A \vee B}$$

$\lor E$

$$\frac{Y(A) \vdash C \quad Y(B) \vdash C \quad X \vdash A \lor B}{Y(X) \vdash C}$$

Next I will define a generic necessity operator based on Greg Restall's [2000] proof-theoretical approach to necessity in his introduction to substructural logics.[92] I will presently show how we can modify this definition to introduce the notion of multiple sources of moral necessities, i.e., a plurality of moral codes. After I prove the Rawls-like claim I mentioned above, I will show how we can safely and constructively abstract away the details of the particular source of a moral obligation. This is important since we want to be able to make sense of the idea that whatever a moral code prescribes is in fact prescribed—even if a different code prescribes something contrary to it.

For generic necessity, I need a new punctuation mark, the unary $\triangle$.

Definition 4.40 (Generic Necessity Introduction and Elimination rules)

$\square I$

$$\frac{\triangle X \vdash A}{X \vdash \square A}$$

$\square E$

$$\frac{X \vdash \square A}{\triangle X \vdash A}$$

With the rules for necessity stipulated this way, box formula and triangle structures are related in a way that parallels the relationship between semi-colon and fusion. As in the case with ';' there are different structural rules we may include to govern the behavior of

---

[92] C.f. [Restall 2000] pgs. 47-52

△. There are no introduction or elimination rules for ';', nor for △. In both cases the behavior of a punctuation mark is governed by structural rules and its role in rules for connectives. Restall [2000] suggests that we think of △ as a sort of unconditional version of ';' (pg. 47-9.) The idea there is to think of necessities as being implied by 'logic', and △ encodes this information, i.e., when it takes a structure as an argument, it entails that whatever follows from that structure is provable by logic alone.

However, a single generic necessity connective is not well suited for a pluralist deontic logic if we need to avoid problematic inferences, i.e., where inconsistent obligations show up. First off, we don't want to be able to infer from A being obligatory and B being obligatory that the conjunction 'A and B' is obligatory. Otherwise, the conjunction of inconsistent actions would be obligatory from the fact of having inconsistent obligations found in pluralist settings. And being obligated to do the impossible is not a welcome result for liberal, i.e., pluralist, foundations. The proof is straightforward, but instructive for sorting out what we will need to define a connective that's a better fit.

Proposition 4.40.1 (Generic Necessity is not well suited for inconsistent obligations)
With a generic necessity connective, a conjunction of necessities entails that the conjunction of the relevant formulas is necessary as well. If we suppose that pluralism permits the conjunction of inconsistent obligations, this would entail that inconsistent actions would be obligatory.

Proof:

$$\frac{\dfrac{\dfrac{\Box A \wedge \Box B \vdash \Box A \wedge \Box B}{\Box A \wedge \Box B \vdash \Box B \qquad \Box A \wedge \Box B \vdash \Box A}}{\dfrac{\triangle\,(\Box A \wedge \Box B) \vdash B \qquad \triangle\,(\Box A \wedge \Box B) \vdash A}{\triangle\,(\Box A \wedge \Box B) \vdash A \wedge B}}}{\Box A \wedge \Box B \vdash \Box(A \wedge B)}$$

And if we suppose that A is incompatible with B we have an unfortunate result.

One might have suspected that a singe generic necessity is not suitable for pluralism if one takes Restall's gloss on the generic necessity at face value. Moral necessities aren't true (or acceptable) come what may or because of 'logic.' I have been supposing that moral obligations are prescribed by moral codes and that pluralism has as its source that whatever a legitimate moral code prescribes is prescribed (together with the claim that there are at least two such moral codes). This suggests that one way to solve the above problem is to instead keep track of the different moral codes that give rise to particular obligations. Instead of a single punctuation mark and necessity, we will need specific punctuation marks corresponding to each of the relevant moral codes. For ease of record keeping, I will use the natural numbers starting with 1 as the punctuation that corresponds to a necessity connective with that number as a subscript. Here are the rules for these connectives.

Definition 4.41 (Multiple Necessities Introduction and Elimination rules)

$\Box_n I$

$$\frac{\mathbb{N}X \vdash A}{X \vdash \Box_n A}$$

$\Box_n E$

$$\frac{X \vdash \Box_n A}{\mathbb{N}X \vdash A}$$

With this connective I block the earlier proof that a conjunction of obligations from two different sources (i.e., moral codes) proves that the conjunction of the topic actions is itself obligatory.

$$\cfrac{\cfrac{\Box_1 A \wedge \Box_2 B \vdash \Box_1 A \wedge \Box_2 B}{\cfrac{\Box_1 A \wedge \Box_2 B \vdash \Box_1 A}{1(\Box_1 A \wedge \Box_2 B) \vdash A} \qquad \cfrac{\Box_1 A \wedge \Box_2 B \vdash \Box_2 B}{2(\Box_1 A \wedge \Box_2 B) \vdash B}}}{}$$

And since the antecedent structures are different, we can't conjoin the respective

consequents, i.e., *A* with *B*, before we introduce any of the necessities.

This modification to necessity not only blocks the derivation of impossible

obligations, it allows me to prove the Rawls-like claim mentioned earlier. But before I do,

let me put the symbols into words. Consider the sequent we will prove below: $\Box_1 A \wedge$

$\Box_2 B \vdash \Box_1 (\Box_2 B \to (A \circ \Box_2 B))$. The consequent may be interpreted as claiming that it is

obligatory by the lights of moral code 1 that if *B* is obligatory because of moral code 2

then choosing *A* is compatible with the obligation to choose *B* from moral code 2. Notice

it doesn't say that choosing *A* is compatible with choosing *B*, only that acting on one code

is compatible with the obligations from another code.

Proposition 4.42 (The moral codes in a plurality must reference the content of each other
in terms of compatibility)

Proof:

$$\cfrac{\cfrac{\cfrac{\cfrac{\Box_1 A \wedge \Box_2 B \vdash \Box_1 A \wedge \Box_2 B}{\Box_1 A \wedge \Box_2 B \vdash \Box_1 A}}{1(\Box_1 A \wedge \Box_2 B) \vdash A} \qquad \Box_2 B \vdash \Box_2 B}{1(\Box_1 A \wedge \Box_2 B); \Box_2 B \vdash A \circ \Box_2 B}}{\cfrac{1(\Box_1 A \wedge \Box_2 B) \vdash \Box_2 B \to (A \circ \Box_2 B)}{\Box_1 A \wedge \Box_2 B \vdash \Box_1 (\Box_2 B \to (A \circ \Box_2 B))}}$$

Theorem 4.42 can be read as a criterion, not for what counts as a moral code

simpliciter, rather for whether a certain group of moral codes are potentially an

acceptable plurality (like the liberal context). As you can see, there is nothing in the theorem that could tell us, if the compatibility doesn't obtain, which moral code is at fault and should be rejected. For example, suppose we have a group of codes that satisfy the criterion from Theorem 4.42, but we then add a new moral code that makes this new group fail the test. There is no formal reason so far that allows us to select the new moral code as the one at fault; in fact, it's possible that the newer moral code merely reveals a hitherto unknown fault with the older group. A good formalism doesn't replace old-fashioned hard philosophical work, after all.

The sequent just proved is related in interesting ways to the familiar—but unprovable—claim that 'ought' implies 'can'. While details regarding how to interpret fusion in terms of a sort of compatibility must wait until some of the formal semantics is in place, this much seems to be entailed by what I have already proved: acting on one moral code means being in a situation where those who hold to conflicting obligations have no reason to object to your choice, or else it's not clear what 'compatible with' could mean. While this superficially resembles Scanlon's [1998: 153] notion of 'there being no reasons to object' as a necessary and sufficient condition for the acceptability of a principle, these two notions are actually quite different. In my case, the claim is much weaker; I have identified a general moral obligation to find at least one situation such that there are no suitable reasons to object to your action, i.e., opposing views don't have to find our choices acceptable or obligatory, nor should we expect that to be possible in general.[93]

---

[93] This fact also feeds into why double negation elimination, i.e., $\sim\sim A \rightarrow A$, is rejected in what follows—more on this below.

There are also interesting connections with some claims I proved in the order-theory, especially the liberal g-graphs of 4.24 and relevant propositions. There I showed that a principled liberal moral code must give up the choice to get their way, suitably construed, even if that option becomes readily available, i.e., without 'need' for compromise. I interpret these claims as all pointing to a sort of trumping obligation in any legitimate moral code (i.e., for a pluralist context): that you ought to find some sort of consensus or compromise as a condition on acting on any one moral code. If this connection bears out, the *compatibility* from 4.42 likely means just this very thing! Thus, if I am correct, the results from the order-theory can provide some guidance for when moral codes are incompatible in terms of 4.42, how to tell which moral code is at fault: the one that claims you ought to get your own way if you can.

It may be useful to present some concrete examples that may be seen as illustrations of the sort of situation addressed by this simple theorem. During the first two years of U.S. President Barack Obama's first term, the Democrats held the balance of power in each legislative body as well as, obviously, the executive. Yet, assuming that the Democrats are generally good liberals and that they made the (arguably false) assumption that their political rivals were working with a different but still liberally acceptable moral code, they should nonetheless have declined the opportunity to get their own way in favor of searching for a consensus or compromise with their political opponents—or, more precisely, to arrive at actions that left their opponents without grounds for objection, i.e., in some sense compatible with their opponents moral code. And it appears they did so (perhaps breaking the hearts of some illiberal leftists….). On the other hand, since Prime Minister Harper's Conservatives won a majority government

175

in Canada, it seems that the only competing values that mattered are those of Canada's Supreme Court. If it's fair to suggest that the Conservatives are generally illiberal[94] then it should not be surprising that as soon as they get their turn to get their own way, they do so, i.e., they seem not to care whether their policies could be compatible with moral codes from opposing views. The basic idea here is that liberals see politics as a way for disagreeing groups to work together, to cooperate politically, socially, and economically. On the other hand, illiberal folk generally see politics as a winner-takes-all proposition.

Thus, insofar as a conjunction of obligations from different moral codes can be construed as representing a pluralist setting for meta-ethics or liberal foundations, this proof demonstrates a Rawls-like claim regarding the content of moral codes in a pluralist setting: a sort of compatibility or tolerance. And this proof does so merely in virtue of the most bare bones meaning of the logical words involved: obligation, conjunction, fusion and implication.

We can also gloss this result in the following way: in a pluralist setting, moral codes must (in a logical sense) take into account the obligations found in other relevant moral codes. And this fact follows merely from the weakest notions of the logical words involved. Of course, it follows that adding stronger rules won't affect this consequence: when we add more structural rules we can prove more, not less.

We are now ready to move on to the general conditions for the semantic frames I shall argue are a good fit for a deontic logic under conditions of pluralism. After we have

---

[94] Of course this line of argument assumes that illiberalism goes along with a rejection of pluralism. While this seems to me to be generally true, I have not argued whether alternative general political stances could give rise to alternative pluralisms.

explored that issue, we can then align the structural rules and axioms with our frame semantics.

4.43 The Frames for PRIT-friendly logics

We turn now to laying out more systematically a formal semantics in which the conceptual tools just described can be embedded, i.e., a formal system suitable for representing liberal pluralism. Such a system will need to be paraconsistent, relevant and constructive: paraconsistent because of disagreement between codes is allowed for, relevant because adding premises to a code will change what is obligatory (as in a strong permission), and constructive because of the nature of the sort of liberalism we are looking at. The system on offer, *RD- choice* (for Deontic Relevant logic of choices[95]), is, I suggest, suitable for investigating obligation under such pluralist conditions.

It is common in introductions to modal logic to suggest that the concept of *obligation* may be represented as a sort of moral necessity. According to this perspective, we model *obligation* as a logical word defined in terms of a class of modal frames. Specifically, an obligation in a world is defined as formulas true at all worlds morally accessible from it. In the order-theory presented above, *obligation* is independently analyzed as a sort of predicate; I will show below that it is a predicate that can be turned back into a logical word. Moreover, there are three further changes to the more familiar formalization of obligation I will investigate.

---

[95] Saying that this system is based on any particular relevant logic, such as *R*, would be misleading. The logic here, while small "*r*" relevant, is much weaker than is *R*; in particular, negation rules are more constructive than generally found in relevant logics.

While obligations and permissions in *RD- choice* will be represented in terms of truth and falsity conditions in the class of relevant models, this is best understood as the acceptability of *plans*, which, as described above, are officially the interpretations assigned to the formulas of the language.

Secondly, we will need a relevant arrow to characterize implication in any deontic logic that avoids a class of validities that would be disastrous in moral discourse. The relevant arrow, together with a relevant and constructive negation I introduce, will also result in a paraconsistent system.

Thirdly, we must adapt the semantics to allow for several distinct sources of obligation. We could do so by allowing for a multitude of accessibility relations, each relation representing a different moral code (for example). However, the same results (with added benefits moreover) can be accomplished with the more elegant technique of using neighborhood functions to represent *necessities*—in our case, moral necessities. A neighborhood function takes a world and returns a class of worlds; the propositions 'true' at these worlds are then the propositions necessary relative to the input world. This technique allows us to transform the predicate-analysis of obligation into the more formally useful *logical word* approach.

The final issue is that while individual moral codes, the ideals of our regulatory ideal theories, will be complete and consistent, how they interact is still a matter of constructive process. In other words, even though we could use classical logic to reason *within* a particular RIT, when we must reason between them, the reasoning must be constructive, paraconsistent and relevant. And as may well be surmised, what can be proved using the rules corresponding to such weak rules is much less than one may have

178

expected or hoped for. On the other hand, as I have already shown, important and surprising truths of liberal foundations can still be proved, even with such weak logical machinery. Moreover, a disappointment can still be an important philosophical lesson.

4.44 Not True and False, Not Acceptable and Objectionable

Talk of *choices* or *plans* being true or false seems awkward and forced. I propose that we instead appeal to the more general notion of *designated* values. *Designated values* are those members of the set of semantic values, *V*, that matter to deciding validity: designated values must be preserved in valid inferences, the most well-known case being *truth* in classical logic. For present purposes, instead of truth I will take the semantic values for choices and plans to be *acceptable* or *not acceptable*, with *acceptable* being our designated value. In this case, valid inferences preserve the acceptability of plans instead of preserving the truth of propositions. Conveniently, while they are, philosophically, different notions from truth or falsity, there are precisely two of them. I will therefore use a version of a two-valued possible worlds semantics. In the two-valued semantics for deontic logic I have in mind, this means that options are interpreted as being either acceptable or not acceptable, but not both, in any particular world.

As noted above, we will not be interpreting negation in terms of mere absence of 'truth.' Negation cannot be represented by a plan merely not being acceptable in a paraconsistent system with just two values (since negation would act 'classically'). So I will represent negation in intensional terms: using *fusion* to represent a compatibility relation. That is, given some world *w*, a plan is negated (i.e., objectionable) not when that plan is not acceptable at *w* but when it is not acceptable at any compatible worlds. With

179

this framework, we have a clear difference between options failing to be acceptable and being objectionable.

However, the standard way in relevance literature to specify the class of compatible worlds for any world verifies the rule of double negation elimination (DNE), and DNE, notoriously, is the *sine qua non* of non-constructive reasoning. Whatever one's intuitions about DNE in general or in moral discourse in particular, a suitably philosophical interpretation of the conditions required to make the compatibility relation verify the non-constructive rules generally valid for negation in relevant logic seem implausible in a pluralist deontic setting.

4.45 DNE Is Problematic in a Pluralist Deontic Setting

Consider two moral worlds, *w* and *v*. In a general sense, *w* is compatible with *v* when nothing that *w* rejects is accepted at *v*. But in order for compatibility to matter logically, i.e., as a way of representing negation, there must be a further relationship between *w* and *v*. In the previous chapter I introduced the notion of the hereditariness of truth, and there is a similar notion at play here. A convenient way to think about this property is in terms of 'earlier' and 'later' states of, for example, a research project, where later states generally have more established truths, but which anyway include all the truths established at any earlier state—the idea is that in a constructive process once a truth is decided (proved, established and so on), it can't be overturned by later developments. To facilitate talk of later stages including truths established at earlier

stages, it will be fruitful to think of worlds, semantic points in a frame, situations and so on, as if they were sets of truths.[96]

From a technical point of view, it is easy to build an inclusion relation into our models for operations like conjunction or disjunction. It is a bit trickier for negation; we need particular conditions to ensure that negation is preserved up the inclusion order as well. Since negation will be spelled out in terms of compatibility, we must ensure that compatibility is constrained in terms of inclusion: if a includes b then if a is compatible with c, b is compatible with c; if c includes b then if a is compatible with c, a is compatible with b. So we can say that a world/stage x *objects* to A just in case for each y compatible with x, y doesn't accept A. We say xCy if x is compatible with y.

In order to verify DNE, it will suffice to guarantee that for any world w there is a world compatible with w, say y, such that for any world z compatible with y, every sentence true at z is also true at w.[97] This condition is problematic however. Since we can assume (safely one might think) that since $\sim\sim A$ follows from A, the reason why we shouldn't accept rejecting some plan at w, i.e., wCy and $y \nvDash \sim A$, is in fact just a matter of $\sim\sim A$ and A being provably equivalent, i.e., since $w \vDash \sim\sim A$ and wCy (and thus $y \nvDash \sim A$) for any z, such that yCz, $z \vDash A$ and according to our DNE conditions, whatever is true at z

___

[96] Please don't take this *ersatz* talk as analysis. It is intended merely to enable us to talk as if one world could include another world, for example, and thus make fruitful use of relevant mathematical objects and relations for representing what our logic can and cannot do.

[97] To show this, suppose $w \vDash \sim\sim A$. We need every y, such that xCy, $y \nvDash \sim A$. If we claim that for any z, such that yCz and thus $z \vDash A$, that w includes z, we have $w \vDash A$ as desired. This part of the argument is adapted from Restall [2000] pg. 261. This condition is also necessary: Suppose that $\sim\sim A \vdash A$, thus for any world w, if $w \vDash \sim\sim A$ then $w \vDash A$. We need that for some world v such that $v \sqsubseteq w$, $v \vDash A$ (even if all this means is that $v=w$.) Since $w \vDash \sim\sim A$, we also need some y such that xCy and $y \nvDash \sim A$. We then need some z such that yCz where $z \vDash A$. But in order to ensure that $v \vDash A$, all we have is that $z \vDash A$ and thus can only set $v=z$ to do so.

is true at *w*. But I have already noted that, taking disagreement and compromise seriously,

we need to distinguish between not having reasons to object to some policy and taking

that policy as acceptable. That is, we may recognize that another moral code is an

appropriate source of guidance in a liberal democracy, yet not take it as one's own such

source of guidance. Thus, one moral code may fail to see some policy as acceptable, but

that thereby, in of itself, gives no grounds to object to that policy in every case. We need

to reject this possibility, i.e., that $\sim\sim A$ being provable and A being provable must always

travel together.

4.46 Moral Necessities

The third basic idea underlying this system is the account of obligation and

permission at a world. The usual way to formalize obligation is with possible worlds talk,

where we use a suitable accessibility relation and quantify over ideal or otherwise

morally appropriate worlds. With this approach, when all worlds *morally* accessible from

some world *w* evaluate some choice, action, plan or proposition (as the case may be) as a

designated value, that choice etc. is obligatory at *w*. The problem is that this approach is

natural only if the modal logic is a normal modal logic. But deontic logic in general

seems like a poor fit for being a normal modal logic, let alone a pluralist deontic logic as

in the current project. Thus, before I explain why the accessibility approach is too

convoluted if non-normal, let me argue why we would want to go non-normal in the first

place.

Let me start with a general problem with any deontic logic being normal. Part of

what makes a modal logic normal is that it verifies the inference of $\vdash \Box A$ from $\vdash A$.

Using such a necessity operator to represent obligation would mean that any truth provable from logic alone would thereby be obligatory. Obviously, this allows us to infer an 'ought' from an 'is' in a suspiciously easy manner, but there are further, more pressing problems. First, perhaps we can't work with a free logic, (we need $A=A$ to hold generally, for example), thus our logic assumes that something exists. This implies that something ought to exist rather than nothing and perhaps we have a strange and far too easy moral realism. The other problems become clear once we consider a pluralist deontic context.

The formulas $\Box A \& \Box B \vdash \Box(A \& B)$ and $\Box(A \rightarrow B) \vdash \Box A \rightarrow \Box B$ are equivalent for characterizing normal K modal logics but, as Jennings, and Schotch [1981] point out, an aggregation principle such as the former formula obliterates an important deontic distinction: not being obligated to do the absurd is different from being permitted to do what one is obligated to do.[98] Leading up to a more concrete example, the latter formula can be glossed in a pluralist deontic context as claiming that if one moral code prescribes that if you plan on doing A then you do B as well entails that if a moral code prescribes plan A then some moral code must also prescribe plan B. But this runs contrary to the whole point of considering a pluralist deontic system in the first place. Say a correct moral code prescribes that if you go to church, you give ten percent of your income to the church. But this code doesn't prescribe going to church, it merely prescribes that if you go, tithe. Perhaps an atheist could even accept this code. But suppose another correct moral code does in fact prescribe going to church but not tithing (perhaps because money could be better spent feeding the hungry). This seems like poor reason for the first code to then go ahead and prescribe tithing to churches or, alternatively, to claim that the latter

---

[98] Proof can also be found in appendix 1.

code is incoherent because it doesn't prescribe tithing or to expect that there must be some other code that does prescribe tithing. But this would be the case if these formulas are valid in a pluralist deontic logic.

At a technical level, non-normal 'accessibility' semantics is a mess to work with. To represent a plurality of moral codes, i.e., with each having its own notion of obligation, is to have each code have its own accessibility relation. An initial problem with such a straightforward formalization of this idea is that a *relevant* version of implication turns out to be both a blessing and a curse. The problem in this case is that the most well understood semantics for the arrow in relevant logics uses a sort of accessibility relation, similar to that found in the possible worlds semantics of modal logics in general. If we have a plurality of accessibility relations, it looks like we would end up with a very messy and perhaps unintelligible structure representing implication (i.e., we would have a plurality of implications and thus a tough time making sense of the corresponding notion of a conditional being appropriately logical).

A simple fix for this problem is to have different sorts of accessibility relations— *moral* and *modal* in this case—after all, it might seem strange to treat implication as somehow *moral* in nature in deontic logic. The problem now is how to fix how the different moral accessibility relations could interact in an intelligible fashion. A quick example of this issue is that if we want to represent *permission* as a sort of deontic diamond, i.e., possibility, operator and define it as usual in accessibility terms, we say that $x \vDash \Diamond A$ *iff* there is some y morally accessible from x and $y \vDash A$. But now consider a dilemma where $x \vDash \Box P$ and $x \vDash \Box Q$ (non-trivially) such that P and Q are incompatible plans. If we define permissibility as above, dilemmas don't seem to really matter any

more since we are permitted to do either, rather than face some hard work sorting which action we should do at the cost of failing to do the other. In other words, at least in some cases, resolving a dilemma means that we are, it turns out, forbidden from doing one action because we must do the other, and letting permission work like this sense of possibility, i.e., be implied by obligation, would undermine this.[99]

There is a fix for this problem as well but it is not going to help us in our project. We could instead [directly] define $\Diamond$ in terms of $\Box$, such that $\Diamond = {\sim}\Box\,{\sim}$. But since we are using compatibility to represent negation conditions, our frames won't let us do this in any elegant, simple, or perhaps even coherent fashion. Consider if $x \vDash {\sim}\Box{\sim}P$ then for each y such that xCy, $y \nvDash \Box{\sim}P$ and this follows with accessibility conditions if and only if there is some z that is morally accessible from y and $z \nvDash {\sim}P$. And this holds if and only if there is some v such that zCv and $v \vDash P$. The point here is what does the fact that there is this one v such that $v \vDash P$ have to do at all with whether $x \vDash \Diamond P$? There is no doubt we would need to ensure that the worlds compatible with worlds accessible from worlds compatible with x are also accessible to or from x. And if that line of reasoning is bit hard to follow, let alone how it could be apt for representing permissibility, you can see why we might want to find some way to go non-normal without recourse to accessibility relations.

The lesson here is that we need a non-normal modal logic to do a pluralist deontic context justice, but if we could find one that handles non-normality without defining our *deontic* modalities in terms of accessibility relations, we would be headed in the right

---

[99] As may be guessed, this line of reasoning means that generally obligation doesn't imply permission in a pluralist deontic setting—see 4.66-8 for when this inference is or isn't valid in PRIT.

direction, or at least a better direction. It is in this light that I propose we instead use *neighborhood semantics*.

Rather than a set of multiple accessibility relations, we can abstract out the mathematical notion underlying such a structure. The idea is that we could take a set of formulas to represent all the plans some correct liberal theory consider obligatory (say, as developed in the graph-theoretic approach from earlier) and because we are assuming a pluralist setting, we will have multiple such sets, one for each correct policy-guiding code. For each world we could then identify a set containing all these sets of formulas,[100] each such subset thus representing what the different liberal theories consider *obligatory* for that world. Thus, I need a mathematical structure that will tell us, for each world, the set of sets of obligatory plans. We call the subsets of these structures the neighborhoods of our input world, and this is why we call the use of such mathematical objects to represent modalities '*neighborhood* semantics.'

To help further motivate the decision to use neighborhood models instead of the more recognizable relational models, I want to point out the many formulas that are not valid on the class of all neighborhood frames. These formulas may all be seen as problematic for deontic logic, and using a technique that invalidates them right off the bat is therefore a virtue of the theory. Without going into too much detail regarding why each of these formulas are problematic for deontic logic, I will point out a couple of fairly well known philosophical issues from the literature.

---

[100] As may be suspected, formalizing this idea in what follows does in fact verify Russell's paradox. Which is really appropriate if you like relevant logics.

In appendix 1 I prove that the following formulas are not valid for the class of neighborhood frames. (I will use $O\psi$ to represent *obligated to choose ψ* and $P\phi$ to represent *weakly permitted to choose ϕ.*)

a. $O\psi \,\&\, O\phi \rightarrow O(\psi\&\phi)$

b. $O(\phi \rightarrow \psi) \rightarrow (O\phi \rightarrow O\psi)$

c. $O\phi \rightarrow \phi$

d. $O(\phi\&\psi) \rightarrow O\phi\&O\psi$

e. $(\phi \rightarrow \psi) \rightarrow (O\phi \rightarrow O\psi)$

I have already argued that (a.) and (b.) don't play nice with a pluralist deontic context. Formula (c.) represents the inference that if anything is obligatory, it is true; since we do not live in a paradise, there are many obligations that are not discharged. Of course, this is readily enough rendered invalid in standard deontic logics by not having accessibility relations that are reflexive.

The schemas (d.) and (e.) are equivalent and are valid in all normal modal logics. That they should be viewed as problematic comes from the fact hat they seem to give rise to things like the Gentle Murder Paradox [Goble 1991]. Recall that the paradox arises if we grant the plausible assumption that, if you kill someone, you ought to kill him or her gently (or in self-defense.) Of course, it's a logical truth that if you kill someone gently, you kill him or her. By an instance of (e.) and detachment, we have that if you ought to kill someone gently, you ought to kill him or her. But this, given simple propositional principles and the aforementioned assumption, gives us that 'if you kill someone, you ought to kill him or her.' If something has to go to resolve this problem, an inference like (e.) is likely less plausible than the principle that one ought to always reduce harm when

one can or generally, to only kill in self-defense. If this is correct, the best way to resolve this paradox is to jettison (e.) and this can be done easily by switching to neighborhood models.

Instead of rigging together relational semantics to avoid these formulas, for each formula at least one structure induced by neighborhood semantics circumvents them—see the appendix for details. It's quite remarkable that as we have philosophical reasons to reject these formulas in the deontic domain, that neighborhood semantics comes ready made to invalidate them on the formal level. Of course, this is just a consequence of neighborhood semantics being a generalization of accessibility semantics (to show this is beyond the scope of this project but we can see how this would work in the one direction, i.e., as evidenced by the fact that neighborhood semantics doesn't validate the normality formulas). It looks, at least as a starting point, that neighborhood semantics is better suited for pluralist deontic logics than is relational semantics, despite the relative unfamiliarity of the neighborhood approach. Hopefully, the fact that neighborhood models generalize over relational models will help offset any uncomfortable novelty.

We have two other advantages from going with neighborhood semantics, i.e., for representing a pluralist deontic context. First, I want to stress how well neighborhood structures fit with our earlier graph-theoretic discussion. With a relational approach, one might see the truth conditions for obligation as a sort of substantial analysis, i.e., obligation here is what is true (acceptable and so on) at any morally ideal (better and so on) world relative to us. But with the neighborhood approach, the neighborhoods for a world are just those sets of formula corresponding to the maximal elements from the relevant graphs, i.e., our moral codes. Of course, one might think that the accessibility

relations would merely encode the same information as found in our graphs without implying anything more philosophically substantial. But the point is that with the neighborhoods approach, reading off our models a competing, substantial analysis of obligation is not even a possibility: the models just are the right sort of mathematical structures to encode the information found in our graphs. In other words, the semantics are formal, not philosophical (or not so in a way that undermines the analysis of obligation found in the graph-theory).

The second advantage is that negation will behave a little more simply when it modifies **O** and **P**. This is a feature of the definitions and neighborhood models. While we evaluate plans as acceptable or objectionable (or both or neither), obligations and permissions are either true or not true of the relevant neighborhoods.

4.47 Frame Semantics

We start with definitions for the objects we will use to build our frames. *Atomic Plans* are what will be represented by the atomic formulas of our formal language. To a first approximation, these *Atomic Plans* will be interpreted by sets of points, i.e., a plan to do *A* will be the set of points where the atomic formula *A* is acceptable—this is essentially the familiar idea of an *atomic proposition* in the semantics for modal logics. However, to interpret non-atomic formulas, we need to generalize this notion somewhat. We therefore introduce another notion, that of *point-sets*. Importantly*, point-sets* have some nice mathematical properties related to graph theory [Cardinal, Hoffmann, Kusters 2013; Brandenburg 2008]. This will facilitate our argument that this formal semantics

189

captures how we should reason about obligation (as found in the order-theory section) all

that much more natural.

Definition 4.48 (Points and point sets)
(Under a particular interpretation) we identify each point with the set of formulas 'true'
there, and write $x \vDash y$ if every formula 'true' at x is also 'true' at y. If P is a non-empty set
of points, the pair $\mathcal{P} = \langle P, \sqsubseteq \rangle$ is a *point set* and $\sqsubseteq$ is a partial order on P.

Definition 4.48.1 (The set of Plans)
The set $\text{Plan}(\mathcal{P})$ of *plans* on a point set is the set of those subsets of P that are $\sqsubseteq$-closed
upwards.

Technically speaking, the plans of 4.48.1 are *order filters* on $\text{Plan}(\mathcal{P})$. If $\{A\}$ is a

point set, then there will be a smallest plan generated by A, *viz*. the set of all point sets

that include $\{A\}$. This will be the official interpretation of the atomic formula A,

replacing the intuitive idea of an *Atomic Plan* described above. The generalization will be

helpful when we look at neighborhood semantics—more on this below. It is worth noting

that, while in the graph theory we used *situations* to represent sets of *choices*, here we use

the more general notion of *point sets, and in particular of plans,* that can represent sets of

formula.

Formulas are connected with more than just the extensional connectives familiar

from introductory logic courses, but also with intensional variants. We turn now to

defining those relations on our point sets suitable for representing these intensional

connectives.

Definition 4.49 (Point Set Relations)
A binary relation C on a point set is a *d-compatibility relation* if and only if for any
$w, x, y, z \in \mathcal{P}$ if $w \sqsubseteq x$ and $y \sqsubseteq z$ whenever xCz then wCy.
A ternary relation R on a point set is an *d-implication relation* if and only if for any
$u, v, w, x, y, z \in \mathcal{P}$ if $u \sqsubseteq x$, $v \sqsubseteq y$, $w \sqsubseteq z$, and $Rxyz$ then $Ruvw$.

Before I can define a frame, I need one more notion. I didn't define this notion for the proof theory since it made no appearance in anything then under discussion. However, we need a way to register *theoremhood* in the logic, i.e., to show what follows from logic alone. To do so in the semantics, we need the notion of an identity for fusion. In the proof-theory, identity for fusion is equivalent to a punctuation mark we can use to indicate theorems. It might be slightly amusing to note that we can define this element without first defining fusion—see 4.52.

Definition 4.50 (Sets apt for identities for R)
We define an identity for the relation R such that for any $T \in Plan(\mathcal{P})$ it is *left identity* if and only if for each $x, y \in \mathcal{P}$, $x \sqsubseteq y$ iff $\exists z \in T$, $Rzxy$ and *right* when $Rxzy$.

We can thus use elements of T to represent that 'whatever is true at x is true at y' is a theorem—more on this below.

Definition 4.51 (D-Frames)
A *d-frame* $\mathcal{F}$ is a point set with the above relations and identity sets.

Definition 4.52 (Evaluations on D-frames without modal connectives)
Given a d-frame and formulas A, B, we say that a relation $\vDash_+$ is an *evaluation* if and only if for each non-modal connective from the proof theory, i.e., without negation, necessity or permission, we have the following conditions on the frame (I drop the subscript where context makes it obvious):
- $w \vDash p$ iff $\forall \mathcal{F}$, there is some $X \subseteq \text{Plan}(\mathcal{F})$ such that $\forall x \in X$, the evaluation assigns p to x and whatever is true at x is true at w
- $w \vDash A \wedge B$ iff $\forall \mathcal{F}$ $w \in \mathcal{F}, w \vDash A$ and $w \vDash B$
- $w \vDash A \vee B$ iff $\forall \mathcal{F}$ $w \in \mathcal{F}, w \vDash A$ or $w \vDash B$
- $w \vDash A \rightarrow B$ iff $\forall \mathcal{F}, \forall w, y, z \in \mathcal{F}$, where $Rwyz$, if $y \vDash A$ then $z \vDash B$.
- $w \vDash A \circ B$ iff $\forall \mathcal{F}, \forall w \in \mathcal{F}$ there is some $y, z \in \mathcal{F}$, where $Ryzw$, $y \vDash A$ and $z \vDash B$
- $w \vDash \bot$ for no $x \in \mathcal{F}$ for any $\mathcal{F}$, i.e., if $x \vDash A$ and $x \nvDash A$ for some arbitrary A, $x \vDash \bot$

The value $[\![A]\!]$ of a formula A according to $\vDash$ is the set $\{w \in \mathcal{F} : w \vDash A\}$ of all those points where A is true or acceptable.

Definition 4.52.1 (D-Models)
A *d-model* $\mathfrak{M}$ is a d-frame with an evaluation on that frame.

I am now in a position to sort out what structural rules I need for the proof-theory and thus the corresponding constraints on R and C relations for the *d-frames*. The main benefit of approaching logical inquiry in this fashion is that it allows us to carefully add in structural rules and conditions on frame relations, one at a time. This allows us to adjust our rules and semantics in relationship to one another, enabling us to ensure that our connectives behave as called for. And if we find the rules and semantic conditions of a philosophically satisfying sort, we might then accept inferences validated by them that we may have thought wrong-headed before. Alternatively, if we dislike the conclusion, we can get a clearer picture of the philosophical costs (in terms of rejected structural rules and awkward semantics) involved in avoiding it. This way logical inquiry can justify its outcomes in terms of finding a best fit between the meaning of connectives and the class of inferences we validate [Goodman 1971].

4.53 Rules and Conditions for Negation and the Ternary Relation.

First, I will need a way to ensure that the R relation in the models line up with the proof-theory structural rules governing ';'.[101]

Definition 4.54 (R-Terms for structural rules)
$R(st)uw = \exists v(Rstv \wedge Rvuw)$
$Rs(tu)w = \exists v(Rtuv \wedge Rsvw)$

These terms will let us encode conditions on d-frames that correspond to structural rules governing ';' once we specify how our evaluation relates points to structures.

---

[101] Again, the formalism here largely follows the presentation in Restall [2000], though I correct for a small misprint for the Associativity rules.

Definition 4.54.1 (Evaluation relating Points and Structures)
$w \vDash X; Y$ *iff* for some $x, y \in \mathcal{F}$, where *Rxyw*, $x \vDash X$ and $y \vDash Y$

This definition clearly matches the idea that formulas connected with fusion are

equivalent to the corresponding structures with ';' punctuation marks. Thus, the '*v*' in

definition 4.54 represents a sort of a proof of, or a designated formula for, the relevant

fusion connected formula.

Just as a great virtue of Kripke semantics for normal modal logics is that key

modal principles at the syntactical level correspond exactly to conditions on the modal

accessibility relation, in substructural logics many structural rules correspond exactly to

conditions of the ternary relation R.

Proposition 4.55 (Structural rules and conditions on R)
Any of the following structural rules correspond exactly to the specified conditions on R.

$X; (Y; Z) \Leftarrow (X; Y); Z$	*Associativity*	$R(st)uw \rightarrow Rs(tu)w$
$X; (Y; Z) \Leftarrow (Y; X); Z$	*Twisted Associativity*	$R(ts)uw \rightarrow Rs(tu)w$
$(X; Y); Z \Leftarrow X; (Y; Z)$	*Converse Associativity*	$Rs(tu)w \rightarrow R(st)uw$
$(X; Y); Z \Leftarrow (X; Z); Y$	*Commutativity*	$R(su)tw \rightarrow R(st)uw$
$X; Y \Leftarrow Y; X$	*Weak Commutativity*	$Rstu \rightarrow Rtsu$
$(X; Y); Y \Leftarrow X; Y$	*Contraction*	$Rstu \rightarrow R(st)tu$
$X; X \Leftarrow X$	*Weak Contraction*	$Rsss$
$X \Leftarrow X; Y$	*Thinning/Weakening*	$Rstu \rightarrow s \sqsubseteq u$

Proof: See Restall [2000] *pgs. 248-52*

Throughout different chapters I have suggested that the order of premises doesn't

have to matter at this level of representation. Hence, the commutativity rules shall be

assumed without further argument.

Since we can think of brackets and punctuation marks as allowing us to classify or

label bodies of information, theory, premises etc. in terms of composing one sort with

another, establishing which structural rules we want can be a matter of arguing how type

composition should work in some domain. For example, whether composing *X's* with *Y's* and then composing this composite with a *Z* is *not* the same as composing an *X* with a composite of *Y* 's with *Z's* may depend on whether composition gives unique outputs, such as we might find when composing, and thus typing, functions. On the other hand, when composition doesn't give such unique results, i.e., when composition of the form (*X;Y*);*Z* results in the same bodies of information or theories (and so on) as composition of the form *X*;(*Y;Z*), the associativity rules are needed. I have argued that issues of the order in which we choose our actions and thus combine our plans is better dealt with at the level of *choices,* not at the more abstract level of *plans*. This means that plan composition is associative.

When is type composition not idempotent? Idempotent composition means that whatever follows from *X;X* follows from *X* on its own. Some logics need to be 'resource-sensitive', i.e., you might need two copies of the same bit of information, not just one, to trigger a state change, for example. Resource-sensitive logics are thus not idempotent. The various linear logics are resource-sensitive and show up when we reason about computability or state machines. Typing and combining plans is not resource-sensitive and thus should instead be idempotent: what follows from two copies of the same plan, follow from just one copy of the same plan.

Classical and intuitionist logic are monotonic in the sense that if *Q* follows from *P*, *Q* will follow from *P* composed with any *X* whatsoever. It is this monotonic feature that verifies irrelevant inferences in such logics, and as argued earlier, we need to ensure such inferences aren't permitted. Thus, composing plans shouldn't be monotonic and therefore we jettison thinning or weakening.

194

With our conditions on the *R* relation set to correspond to these structural rules,

we can move on to model conditions for negation and their corresponding proof rules.

Definition 4.56 (Negation)
- $w \vDash \neg A$ *iff* for some y such that wCy, $y \nvDash A$

With this condition for negation, the *C* relation says nothing whether y is compatible with

x. This brings up a possibility to define another sort of negation where the compatibility

relation takes the arguments in the other direction, i.e., not *A* at *x* when yCx and y doesn't

accept *A*. Arguably, there is no need for these two sorts of negation. To avoid this, we can

set the *C* relation as symmetrical. The next issue is how fusion, implication and negation

are interrelated. First off, if we have need to, we should be able to represent negation in

terms of 'implying the absurd' (or something similar). Second, I have previously helped

myself to interpreting fusion as a sort of compatibility; I can now live up to my promises.

And thirdly, we have already seen how fusion and implication are related from the

formalism so far. But before we get to the conditions for frames, let me introduce the

corresponding rule for negation in our proof-theory.

Definition 4.56.1 (Rule for Semicolon Negation)
$\neg I; \neg E$:

$$\frac{X; A \vdash \neg B \qquad Y \vdash B}{X; Y \vdash \neg A}$$

Notice that this one rule both eliminates a negation from a formula in a premise and

introduces a negated formula in the consequent.

Definition 4.56.2 (Frame Conditions for Negation)

$xCy \rightarrow yCx$	*Unique Negation*
$\exists x(Ryzx \wedge xCw) \leftrightarrow \exists v(Rywv \wedge zCv)$	$\neg I; \neg E$
$\forall x \exists y\, (xCy)$	$\forall z \in T, \neg z \vDash \perp$

195

The frame condition for $\neg I; \neg E$ looks fairly daunting. It helps, I think, to consider the

proof that these are precisely the conditions we need to make the negation rules valid.

Proof: In the first direction, we are looking to ensure if $C; A \vdash \neg B$ then $C; B \vdash \neg A$. This means we need to know, supposing that $C; A \vdash \neg B$, that if $x \vDash C; B$ then $x \vDash \neg A$. Thus, we need to see that there is some w such that xCw and $w \nvDash A$. Suppose that $x \vDash C; B$. Next, since $Ryzx$, there is a y, z such that $y \vDash C$ and $z \vDash B$ by the definition of fusion. We also know that $\exists v(Rywv \wedge zCv)$. Thus, $v \nvDash \neg B$ because zCv and $z \vDash B$ from earlier. But now, since $C; A \vdash \neg B$, this implies that $v \nvDash C; A$. And with having a y, w such that $Rywv$, we have $y \vDash C$ from earlier. This implies $w \nvDash A$, as required. The other direction works in the same way.

And since fusion and implication are defined as they are in 4.52, this condition

ensures that the compatibility relation maintains the right sort of connection with the

implication relation. The third negation condition is just what we expect if elements of T

are supposed to register theorems.

I can now move on to define validity for our system, but first I need to define the

sort of points or worlds that matter for determining validities.

Definition 4.57 (Worlds and Non-normal points)
Non-normal points are any $x \in \mathcal{F}$ such that for some A, $x \vDash A$ and $x \vDash \neg A$ (i.e., it is impossible for xCx) or that for some A, $x \nvDash A \vee \neg A$. Any other such x is a *world, i.e., a normal* point. For any element of the identities for R, i.e., T (4.50), say z, whatever is true at z is true at any normal point.

The idea here is that it is possible that not all points are the same sort of thing in frame

semantics. While we may start in the same place as the more familiar logics, i.e., looking

at what relationships we find between worlds, we extend this idea to include the notion

that worlds may also be related to non-world things, such as incomplete or unrealizable

states of affairs. Intuitively, in a pluralist deontic setting we may have reason to use

unrealizable plans or incomplete information in deductions. Having worlds relate to

incomplete or unrealizable points in a frame allows us to represent how we should reason

under such conditions. And just like the more familiar modal logics, validity is a matter

of what 'truths' hold at every world. We can use the identities for R to do this, as laid out

in 4.50, because of two features. First, the truths of any identity for R are true at any

world. Secondly, in virtue of how identities are defined, we don't have to worry about

there being truths at every world that identities can't register as a theorem.

Definition 4.58 (Validity)
P is a *valid* formula in $\mathcal{M}$ *iff* $\forall \mathcal{F}\ \forall z \in T, z \vDash P$.

As required, any time, i.e., in any frame, an element from our identities for R verifies

some formula, that formula is valid.

Proposition 4.59 (The logic *RD-choice* is relevant)
Proof: I show the relevance of *RD-choice* by demonstrating that for any two formulas, A, C if $A \rightarrow C$ is a valid inference in *RD-choice*, A and C must share at least one atom. Consider the case where we have two worlds, w, u, and a non-normal point x such wCu and that $Rwxw$ and $Ruxu$. We suppose that x is a situation where all *choices* are designated. There are two cases to be considered: (1) when all of A's atoms are only designated at w and all of C's atoms are only designated at u (2) when all of A's atoms are only designated at u and all of C's atoms are only designated at w. [This guarantees that A and C will never share an atom].
Case (1) take $Rwxw$ and suppose that $w \vDash A \rightarrow C$. Thus since $x \vDash A$, we need $w \vDash C$. But in this case C's atoms are not designated at w and thus $w \nvDash A \rightarrow C$ contrary to our assumption (and because of this invalidity there is no need to cover $Ruxu$ for case (1)).
Case (2) take $Ruxu$ and suppose that $u \vDash A \rightarrow C$. Thus since $x \vDash A$, we need $u \vDash C$ But in this case, C's atoms are not designated at u and thus $u \nvDash A \rightarrow C$ contrary to our assumption (and again, one invalidity is sufficient, so there is no need to cover $Rwxw$ for case (2)).

Proposition 4.59.1 (The logic *RD-choice* is paraconsistent.)
Proof: I show paraconsistency in virtue of negation failing to obey the rule of explosion. Explosion is when $p \wedge \neg p \vDash q$ for any arbitrary q. The negation condition on our frame needed to get explosion is that xCx for each x in our frame. Here's how: suppose the explosion condition holds, and that $x \vDash A \wedge \neg A$. Since xCx, $x \nvDash \neg A$ because $x \vDash A$ and thus $A \wedge \neg A \vdash \bot$. Therefore if $x \vDash A \wedge \neg A, x \vDash \bot$ and thus $x \vDash B$ for any arbitrary B. In the other direction, suppose $p \wedge \neg p \vDash \bot$ is valid but that xCx doesn't hold. Suppose that $x \vDash A \wedge \neg A$ and thus everything is true and false at x. But since it also follows that

$x \nvDash \neg B$ and $x \nvDash B$ for any arbitrary B, x is compatible with itself, contrary to our assumption. The condition necessary for explosion is not included in the semantics.

Proposition 4.59.2 (The logic *RD-choice* is constructive.)
Proof: We have already seen that our frames don't have the conditions necessary for DNE. The remaining condition is that excluded middle doesn't hold. The excluded middle is that for any point $x \in \mathcal{F}, x \vDash A \vee \neg A$ for any arbitrary A. The frame condition: for each $x, y \in \mathcal{F}, \forall x \exists y (xCy \wedge (xCy \rightarrow y \sqsubseteq x))$; for sufficiency, suppose that the frame condition holds. If $x \nvDash A$ and $x \nvDash \neg A$ there is some y s.t. xCy and $y \sqsubseteq x$. Thus $y \nvDash A$ and $y \nvDash \neg A$. But then $x \vDash \neg A$ by negation conditions and x isn't a point in $\mathcal{F}$ after all, i.e., $x \vDash \bot$.

For necessity, suppose that for any point x, $x \vDash A \vee \neg A$ for any arbitrary A but that for some $x, y \in \mathcal{F}, xCy$ and that $y \sqsubseteq x$ fails, i.e., there is some formula true at y but not x, say, P. Since xCy, we have $x \nvDash \neg P$. But then $x \vDash P$ follows according to our supposition and thus $x \vDash \bot$. Our frame conditions (for negation) don't include this condition necessary for excluded middle and thus the law of excluded middle is impossible to verify in our frames.

4.60 Deontic Semantics for *R-choice*: R^D-*choice*

I now introduce a version of neighborhood semantics. Let me start with introducing the

notion of sets paired with a distinguished collection of subsets.

Definition 4.61 Given W is a non-empty set, $_\varphi W$ is the collection of all subsets of W. The pair $\langle W, F \rangle$ where $F \in {}_{\varphi\varphi}W$ i.e., $F \subseteq {}_\varphi W$, F is a set of subsets of W.

I now need a way to assign these subsets to points, each subset being those

formulas obligatory for the point they are assigned to.

Definition 4.62 A *neighborhood function* is an N such that $N: W \rightarrow {}_{\varphi\varphi}W$

N is a function taking each point to a set of sets of points, each of those sets is a

neighborhood of the input point.

Definition 4.62.1 A *neighborhood frame*, F, is a pair $\mathbf{F} = \langle W, N \rangle$ where W is a non-empty set and N is a neighborhood function.

Definition 4.62.2 A neighborhood model is a tuple $\mathcal{M} = \langle \mathbf{F}, I \rangle$, where $\mathbf{F}$ is a neighborhood frame and I is the same interpretation from before (4.52.) Taking our frame

to be characterized as above, *I* is extended to an assignment of semantic values for all formulas as before with the additional condition:

1. $x \vDash^I \boldsymbol{O}A$ iff $[\![A]\!]_I \in N(x)$

Intuitively, an evaluation now also specifies that if a set of formulas is assigned to a point, that formula is obligatory at that point. But not all points are equal; some are worlds, some are incomplete or unrealizable states of affairs. In order to ensure we can distinguish between inconsistent obligations within the same moral codes (which is disastrous) and inconsistent obligations abstracted out of equally correct but conflicting moral codes, we need to restrict permissibility conditions to worlds only. We will see why in the following section.

Definition 4.62.3 (*Permissibility at Worlds*)
Given a $\mathcal{M}$, *I* is also extended to assign semantic values as before, with the additional condition for all $s \in Worlds$

2. $s \vDash^I \boldsymbol{P}A$ iff $W-[\![A]\!]_I \notin N(s)$

If according to an evaluation, a formula isn't forbidden (i.e., the set complement isn't a neighborhood) at a world, it is permitted.

In order to use the neighborhood function to pick out obligations, i.e. a specific sort of necessity, we will need some semantic postulates and other, more general, algebraic properties governing *N*.

4.63 Algebraic Properties of *N*

In order to represent the information found in our moral codes (i.e., g-graphs) the neighborhoods *N* sends a point to will need certain constraints. First, *N* is appropriate for *I* only if no normal point, i.e., world, *w* includes ⊥ as a neighborhood. Next, we need a

superset relation such that any set, Q, that includes a neighborhood for x, Q is also a neighborhood for x. This entails that for every $s \in W$, $W \in N(s)$. Intuitively, the superset relation ensures that our neighborhoods capture the order found in our moral codes.

We also need to have some notion of **O 'not A'** and **P 'not A'** as not merely representing that it is obligatory (or permissible as the case may be) to object to A, but rather that it is obligatory (or permissible as the case may be) to not choose A. Again, we will use x to denote any point and w to denote worlds.

Definition 4.64 (Negation as *Not Choosing*)
  3. $x \vDash^I \boldsymbol{O}{\sim}A$ iff $I(\neg A) \in N(x)$
  4. $w \vDash^I \boldsymbol{P}{\sim}A$ iff $W{-}I(\neg A) \notin N(w)$
  5. $x \vDash^I {\sim}\boldsymbol{O}A$ iff $I(A) \notin N(t)$, where sCt
  6. $w \vDash^I {\sim}\boldsymbol{P}A$ iff $W{-}I(A) \in N(t)$, where sCt

Intuitively, I is telling us that when the formulas obligatory from 4.62.2 on 'closer inspection' include negation operators, we can interpret such obligations now as being obligatory to not choose, rather than as before, obligatory to object. Thus, we can use $\boldsymbol{O}{\sim}A$ to represent it is obligatory to not choose A. In condition (5.) we use the compatibility relation to define a negated obligation at x not when that obligation is not a neighborhood of x, but rather when that obligation is not a neighborhood of some point compatible with x. This ensures that this negation, as it is evaluated in terms of compatibility, still triggers the non-normality conditions underlying 4.57 such that if $x \vDash \boldsymbol{O}A$ and $x \vDash {\sim}\boldsymbol{O}A$, x is non-normal, i.e., that x could not be compatible with itself: an inconsistent state of affairs.

The two different notions of negation here allow us to distinguish between obligations to object to some plan versus an obligation to not choose some plan. After all,

we might have a point at which a plan is both acceptable and objectionable, so it's possible that we could have an obligation to object to that plan, but still be obligated to choose it. For instance, we might be faced with a horrible dilemma, and part of our moral code is that we ought to object to having to do something that forces us to give up doing something we should otherwise do. Nevertheless, we ought to actually choose that objectionable plan (or not as the case may be). Again, the permissibility conditions are restricted to normal points or worlds, evaluating whether an objectionable plan is forbidden (4.) and in the other case (6.) whether a neighborhood of a compatible point forbids the plan.

Besides the general call for making such distinctions, there is a significant philosophical point that this formalism allows us to represent (in fact, I take this point to justify the subtitle for this project, "On why Liberalism must be inconsistent to be correct"). Normal worlds are allowed to consider, as it were, what's going on in incomplete and inconsistent places, not only with respect to how the R relation works, but also in terms of what points can show up in a neighborhood for a world. A normal world w could verify a particular sort of dilemma: $w \vDash OA$ and $w \vDash O{\sim}A$. The conditions for this dilemma, i.e., that both $I(A)$ and $I(\neg A) \in N(w)$, doesn't imply that w is not compatible with itself. And informally, obligation has been analyzed in terms of a best choice, given constraints on social roles, so it shouldn't be surprising that choosing some plan and not choosing some plan could both be best choices for some world. Thus, a normal world could have inconsistent obligations, of a sort, without itself being inconsistent. Of course, acting on both choices isn't going to work, but there is nothing incoherent about the inconsistent action guidance. And of course, we have already seen how in PRIT we

jettison the idea that if we are obligated to do two plans that we are obligated to do the conjunction of those plans.

But this is where we start cooking with gas. As we will see in 4.68, our actual liberal context doesn't meet this condition for being a *world*, i.e., self-compatible relative to deontic consequence. If this is correct, it means that in order for liberalism to have a foundation, the political liberal context must itself be inconsistent. On this analysis, liberalism is an unrealizable plan (or rather is composed of plans that together are unrealizable) and this is the bar our attempts at justification must reach in its foundations. I turn now to further parsing the difference between worlds and points, relative to obligatory and permissible plans.

4.65 Properties of the Neighborhoods

I need to also consider whether R^D-*choice* should have an axiom $OA \to PA$. In relational semantics, this formula corresponds to the *seriality* property: for any point, x, there is a point, y, such that y is accessible from x. For my system, the required semantic postulate is even simpler.

Definition 4.66 (Seriality) A *serial* constraint on a neighborhood function for normal worlds is: $s \vDash^I OA$ then $\bot \notin N(s)$ for all normal points.

We already have this condition from 4.63. No world, i.e., normal point, verifies an obligation to do the absurd. 4.66 merely makes this connection to seriality explicit.

Proposition 4.67 Neighborhoods are serial *iff $OA \to PA$* is a valid formula

Proof: Suppose neighborhoods for worlds are serial but $OA \to PA$ is not valid. Thus $\bot \notin N(w)$, $I_D(A) \in N(w)$ but also $W - I_D(A) \in N(w)$. But then $\bot \in N(w)$. Now suppose

$OA \rightarrow PA$ is valid but neighborhoods aren't serial. So we assume that $w \vDash OA$ and that $\bot \in N(w)$. But now $w \nvDash PA$ since $[\![A]\!] \in N(w)$ for every A.

Seriality doesn't apply to neighborhoods for non-normal points because we need to allow for another sort of dilemma to be true at such points. The type of dilemma that we need to use non-normal points to represent is when we are both obligated to choose some plan and not obligated to choose the same plan. In this sort of case, the point x can't be compatible with itself and is thus a non-normal point. But the seriality rule, if general would require permissibility to be recursive at x and this result would allow an irrelevancy being verified, i.e., if x's were R related to worlds, we could verify: if you are not obligated to do A, it is permitted to do A. Thus, the seriality rule only governs normal choice situations. But of course, validity is only a matter of what formulas are 'true' at normal worlds, so there is no loss here.

The above comment is part of a more general fact regarding non-normal situations. If some point is non-normal then we don't use the recursive procedure for determining some of its semantic values. Rather the model assigns non-normal points into the powerset of some formulas, not just atoms, as is befitting points where the *rules of logic don't apply*. In other semantics for relevant logics, it is generally only formulas with *implication* that doesn't use recursive procedures at non-normal worlds. This is often seen as intuitively plausible because of the close connection between implication and the corresponding consequence relation. However, in RD-choice, this notion of the 'rules of logic' i.e., the consequence relation, is extended to *permission* connectives; in this case, it is because we must take dilemmas seriously and thus, in some sense, open-ended.

As it stands, there is only one axiom for obligation connectives in *RD-choice*.

Separation $O(P \wedge Q) \rightarrow OP \wedge OQ$

Separation follows trivially from neighborhoods being closed under the superset relation.

I now turn to putting some of this formalism to use and explicating how and why pluralist deontic settings such as our 'world' we need to make a distinction between not being obligated to not choose some plan and being permitted to choose some plan. In fact, according to *RD-choice* very little deontic reasoning in settings like ours will tells us at all what is permissible, rather only what we aren't obligated to do, all else being equal. Informally, in a pluralist deontic setting like liberalism, not only do we allow for two equally correct moral codes prescribe conflicting plans, but also that one code may prescribe a plan the other doesn't. If the former sort of disagreement were the only to occur, we could represent a liberal setting in terms of normal worlds. But we also seem to be stuck with acceptable codes both prescribing and failing to prescribe the same plan; thus we must be, deontically speaking, in a point that isn't self-compatible.

Proposition 4.68 The actual world is a non-normal point given the existence of moral dilemmas such that we are both obligated and not obligated to choose some plan.

Proof: *Since we are assuming that liberalism presupposes a plurality of moral and political codes, some of which will say we are obligated to choose A and others say we are not obligated to choose A, the conditions for our setting to be non-normal are met.*

The non-normality of the actual world can also be used to support the idea that an enumeration of moral properties in an actual situation cannot always be used to recursively determine what one ought to or may do.

Proposition 4.69 If we restrict attention to models that include only normal worlds *obligation* is interdefinable with weak permission.

Proof: We need to show that $\boldsymbol{O}A \leftrightarrow \sim\boldsymbol{P}\sim A$ is valid across all neighborhood frames that include only normal points. First, let's suppose that $\boldsymbol{O}A$ is not true at s but that $s \vDash \sim\boldsymbol{P}\sim A$, i.e. a counterexample. Thus $W - [\![\neg A]\!] \in N(t)$, where sCt and so $[\![A]\!] \in N(t)$ since t is also normal. But since $s \nvDash \boldsymbol{O}A$, it follows that $t \vDash \sim\boldsymbol{O}A$. Thus $[\![A]\!] \notin N(t)$, a contradiction. In the other direction, suppose that $s \nvDash \sim\boldsymbol{P}\sim A$ and thus $W - [\![\neg A]\!] \notin N(t)$, where sCt. Assume that $s \vDash \boldsymbol{O}A$, i.e., a counterexample. Since s is normal, $s \nvDash \sim\boldsymbol{O}A$ also. But then $t \vDash \boldsymbol{O}A$ and $[\![A]\!] \in N(t)$, a contradiction.

Proposition 4.69 illuminates the relationship between *must* and *only*, making clear that it parallels the relationship between *best* and *ought*. Letting $\sim\boldsymbol{P}\sim A$ represent the idea that it is impermissible to not choose A, this provides a good characterization of what we must choose, namely A. But since $\boldsymbol{O}A \leftrightarrow \sim\boldsymbol{P}\sim A$ is valid only when limited to normal worlds, this interdefinability only happens when there is only one consistent moral code in play. Thus, if we must choose A, it is because it is the only obligation possible. Therefore, *must* is related to *ought* as *only* is related to *best* in PRIT, thanks to *RD-choice*. It just requires that the political world we are most familiar with turn out to be a morally inconsistent one.

4.70 Summary

In this chapter, I analyzed *obligation* in terms of a special sort of predicate. Using an order-theoretic approach, I showed how we could represent obligation as a best acceptable choice from among alternatives. As a crucial tool for carrying out this analysis, we represented normative codes, moral and political, as graphs, and I looked at various ways to manipulate the graphs to represent different sorts of reasoning. I looked at how we could sort what is best to choose from when faced with only a bunch of bad options. I also found a method for representing what to do if our best option becomes unavailable. These last few techniques proved useful in demonstrating some powerful constraints on

what sort of moral or political codes can count as legitimate or principled in a liberal society.

However, treating obligation as a best choice of acceptable options in its full graphic detail is problematic; that is, if we want to use a formal language to reason about and with these mathematical objects. There are good reasons to abstract away much of the choice structure when we want to use logical connectives. Thus, I turned to the idea of *plans*. After a brief introduction of a weakest proof-theory i.e., with no structural rules, that could represent multiple sources of obligation, I derived a proof that could be characterized as a criterion for what counts as a legitimate moral code under conditions of pluralism. That is, moral codes in a pluralist context must contain an obligation calling for us to find a way so that other moral codes couldn't object to our choices.

I then worked out the various properties fusion and implication should have by introducing a frame semantics. The logic was then proved to be paraconsistent, relevant, and constructive.

Finally, I introduced the technique for taking the predicate analysis of obligation, i.e., from the graph-theory, and representing obligation as a logical word: neighborhood semantics. With the neighborhood function, the logic gathers all the best of acceptable choices relative to an input world or point, and returns with *obligations* corresponding to each such best choice. With this technique, we have a way of treating obligation *as best of acceptable choices* as a logical word without being itself in some sense an analysis of obligation (as, arguably, there is with the standard sort of relational possible world semantics). After explaining some of the ways we need the neighborhood function to behave, I looked at the *seriality* condition for neighborhoods. The seriality condition

would ensure that we could always infer that it is permissible to do what we are obligated to do. With this condition, together with a distinction between different sorts of dilemmas, I proved in a formal setting what had been argued earlier in informal terms: that since there are both types of moral dilemmas, ours is in fact a non-normal 'world.' This gave precise bounds to what it would take to justify liberalism: that liberalism itself is an unrealizable plan, not that it just takes the unrealizable composition of plans seriously. This result also led to my final proof: that there could be only one source of obligations, i.e., no disagreement or conflict could be possible, if we must do something. Thus, the PRIT system can represent the claim that *must* is related to *ought* as *only* is related to *best*. This result means that pluralism doesn't require rejecting a systematic approach. It does however require that if liberalism is cogent, it must be an inconsistent system.

Chapter Five: *Final Remarks and Directions for Future Research*

5.0 Summary

This project has aimed at providing an analysis of obligation and its cognates that can make sense in a pluralist context. I motivated this research by arguing that a pluralist approach to liberal foundations makes sense on two fronts. First, I argued that by allowing for multiple correct, but conflicting, answers to foundational questions, some of the problems facing a systematic foundation for liberalism could be solved. Second, I argued that any constructive approach to the foundations of liberalism is also, *prima facie*, a pluralist approach. But without a clear sense of what *obligation* can mean when there are conflicting principles or obligations, we can't know what sort of problems we may be unleashing upon unsuspecting political theorists.

In order to explain the nature of obligation when we assume pluralism, I argued that we should use some of the ready-made tools used by logicians. Specifically, relevant logicians have been investigating consequence relations that could distinguish between explosive inconsistencies and inconsistencies that can nonetheless be fruitfully used in deductive reasoning. A connected reason for taking a formal approach is that we have a tradition in deontic logic for treating *obligation* as a logical word. It seemed like a good fit for exploring the relevant conceptual possibilities, given that we could find a logic that satisfies our criteria.

We next looked at the nature of ethical or political pluralism in more detail. After all, the clearer the picture of pluralism, the clearer the criteria my analysis might satisfy. In this section I explained and developed what it meant for a multiplicity of theories to be a philosophically relevant pluralism: the pluralism must be non-trivial and have

interesting disagreement between its theories. To do so, I introduced the concept of theories being *rivals*. Rivals may equivocate upon certain properties at some levels of theoretical description, but the conclusions they draw on our target level must use an unequivocal sense of the same predicate. Thus rivals may disagree about how to answer the *same* well-posed question.

Formalizing obligation quickly becomes complicated when we aim to represent a wide range of essential and acceptable inferences found in actual best practices. In this project, I focused on the issue of the order in which we live up to prescriptions. As shown, it is difficult to model how one obligation could trump or override another, especially when standard formalization suggests we treat obligations as propositions or actions true at all morally accessible worlds. But investigating the formal conditions behind trumping isn't merely an exercise in trying to solve puzzles in deontic logic. I argued that principled cooperation and coordination between disagreeing political viewpoints calls for such moral codes to contain an obligation 'to get along' that trumps their other prescriptions, that is, when needed.

Informally, I argued that we should analyze obligations in terms of contrastive principles. I developed this conception of obligations by adapting from the notion of contrastive explanation, as proposed by Van Fraassen. By taking moral principles as answers to a *May I?* question, where the topic and contrast class are specified, as in the explanation case, we can account for the fact that decisions are about choices. By using the account of contrastive principles, we could represent trumping in terms of changing the nature of our action-guiding question. Moreover, the nature of contrastive principles

suggested a powerful analysis of obligation: that we should choose the best acceptable option from among the relevant alternatives.

Finally, the previously informal or syntactical concepts of choices and obligations, specifically the resulting ideality ordering and related moral orderings are treated formally with graph-theory. A bare-bones proof-theory was introduced to reason about and with the concepts defined by this order theory. I then used this formally minimalist language to deduce a criterion for any moral code under conditions of pluralism: that any such code must guarantee that acting upon its prescriptions is compatible with the obligations of rival codes. I argued that this criterion has significant philosophical value when it comes to identifying what sort of moral or political codes could count as deserving a hearing, i.e., in a liberal democracy.

The formal semantics for the language *RD-choice* was then introduced and shown to be paraconsistent, relevant and constructive. Next I introduced a semantics that allows us to ignore the theoretical source of a prescription, as needed for a pluralist version of obligation: neighborhood semantics. And as promised, I proved that 'ought' doesn't imply 'can' let alone 'may' when moral pluralism is assumed. Perhaps more surprising, I proved that our world, assuming liberalism and thus pluralism, is not a normal world in the logical sense, i.e., in terms of deontic logic. This means that what we can say follows from a choice being obligatory is very limited indeed.

Obviously, we can't assume that there is always an answer for any moral question, let alone one unique answer. But the results of this project allow me to say a bit more than that. We can't assume that a rule for cooperating or coordinating being fair today implies that it will be fair tomorrow. They could be trumped after all. We have to treat

what is fair or just, even in terms of fundamental policy, as always being subject to revision, or even in some areas in a case-by-case process. Thus, PRIT promises a powerful way to justify making constitutional law amendable or to even include notwithstanding clauses.

But along with such clear pluralist limitations to the reach of deontic reasoning, I proved some significant claims regarding the nature of moral or political codes. We saw that principled liberals won't take an opportunity to get their own way and that we should find ways so that those who disagree with us can't object to our choices. We also saw that fair coordination matters just as much as fair cooperation. That is, when it comes to theories of how we should live together (while still disagreeing about how to live together and about what is good in life), we need to make sure how we settle on a path forward is fair, not just where we are headed.

5.1 Future Directions

5.1.1 Realism/Anti-realism Debates

In the course of investigating obligation under conditions of pluralism, I have abstracted out some general principles or questions of wider theoretical interest. In Chapter One I showed that constructive foundations in general are also *prima facie* pluralisms. This seems to suggest the following conjecture: that debates regarding the use of logical principles to frame realist/anti-realist questions are far more fraught, to say the least, than may have first been thought. For example, supposing a standard way to go

211

semantic anti-realist in some domain is *via* constructivism, one is thereby a dialetheist if we take this well-trodden path for truth predicates, at least *prima facie*.

5.1.2 Pluralism and Modalities

I introduced here a new technique (or rather synthesized two prior ones) such that that we can analyze some logical words, i.e., modalities, in terms of a class of predicates and the relevant properties of a suitable function. This suggests that what we call 'neighborhood' semantics may have more philosophical value than just the nice mathematical features of being a more general account of possible world semantics.[102] Previously, neighborhood semantics accounts relied on the mathematical properties of the neighborhood function to provide a merely formal analysis of the relevant necessity. If I am correct regarding moral necessity, i.e., obligation, we may have a new philosophical method for identifying and distinguishing between different sorts of necessities: the relevant classes of worlds are predicatively defined in terms of an independent characteristic. This leads to a second question, supposing that many central concepts in philosophy can be analyzed independently of a formal notion of necessity: Why is it that when we can treat certain words as a modality, such as knowledge, obligation, proof etc. that we can then have a pluralism regarding the corresponding subject matter if we use neighborhood semantics to represent them as necessities?

To put the philosophical point more directly: in the examples here (knowledge, obligation etc.) the relevant terms admit of analysis independent of the semantics of

---

[102] That is, with the right properties we can have a function that for any class of Kripke frames and models we get a class of neighborhood models that validates all and only those formulas verified by the Kripke models (but not *vice versa*).

Kripke frames and its cognates, such as possible worlds. When we can combine such an independent analysis with truth-conditions for a type of necessity, a philosophically significant pluralism regarding the concept is possible. And that's surprising.

5.1.3 Philosophy of Language

The paraconsistency (and other logical features) needed to make sense of pluralism have repercussions regarding what the corresponding concepts can even mean. In Chapter Two, I showed that we have a clear principle dividing relativism from pluralism. If X according to theory Y^n is A_n we have relativism. If X according to theory Y^n is A we have pluralism. Pluralism abstracts away or ignores the source, i.e. the relevant theory, of the predication. The idea here is that with pluralism, it doesn't matter to what the terms mean which theory was used. This suggests that some notions of inferential semantics or proof-theoretic accounts of meaning, i.e., doctrines that meaning is derived or read off the *inferences* we can make with the terms, are (at least) incomplete with regards to pluralism.

At one level of analysis we keep track of which theory says what as befitting paraconsistent inferences. With this level of analysis, valid inferences directly depend on parameters corresponding to the relevant theories. But at more abstract levels (i.e., those suitable for pluralist contexts) valid inferences are independent of which theory was originally used. The question here is: how can we know which class of inferences matter for the meaning of the relevant terms? Or if both levels matter, how can we make sense of what the relevant terms mean?

5.1.4 Deontic Logic

Much of the motivation for contrastive principles in Chapter Three came from dissatisfaction with how order puzzles are treated in deontic logic literature. It seems that even other independently invented accounts that purport to use contrastive representations of obligation don't solve these puzzles as clearly as I think is possible. My approach treats the modality as binding over the entire choice structure and I think that this feature will make a large difference. Specifically, I propose that the graph-theory in Chapter Four can be used to represent reasoning that avoids the order puzzles common to deontic modalities and do so in a principled manner.

5.1.5 A New Family of Liberalisms

I propose that by using the liberal abstraction norm of *responsibility* from Chapter One, I could develop a new family of liberalisms. The idea is that we could use the responsibilities of an ordinary fiduciary as a starting point for articulating the concept of a *social fiduciary*. If I am correct, the responsibilities of a social fiduciary could account for treating equivalent concerns equally— with liberty being of paramount concern.

There are three components to how I plan on formulating the social fiduciary perspective. First is in terms of choosing under conditions of ignorance. Here I plan to use the well-trodden metaphor of the veil of ignorance, but I propose a better rule of choice under ignorance than may have been used: the lottery assumption rule. It states that it is never rational to assume that you are in a minority group when choosing under conditions of ignorance. The resulting character of such choices under ignorance forces

us to make explicit and justify a social ontology for our social fiduciaries, i.e., on the pain of very racist, classist, hetero-normative policies being allowed in. It then seems that the responsibilities of a relevant social fiduciary should include minimizing paternalist intervention. I propose that this responsibility, i.e., to minimize paternalist intervention, is what social fiduciary foundations can use to analyze, explain, or represent the value of *political liberty*.

If correct about how this will work out, this approach to liberal foundations would make explicit the role of social ontology when we articulate what makes fundamental policy just and liberal. Moreover, from some preliminary investigations, and despite the seeming tilt towards preferring those who start with more to invest; building a social fiduciary perspective based on a *responsibility* norm actually could end up very egalitarian and perhaps more so than wellbeing based approaches.

5.1.6 Legitimacy of a Liberal State

Given some uncontroversial assumptions, Sen's redundancy complaint from Chapter One seems to underestimate the connection between what I propose is a thin notion of legitimacy, reasonable expectations, and the process of justifying basic institutions. General principles of justice (to adapt a phrase from Nelson Goodman [1983]) commit us to deciding unrealized cases of right. This suggests that questions of legitimacy specifically focus on what injustices the state will remedy. It seems plausible that the notion of legitimacy, in its thinnest form, involves our reasonable expectations regarding the sorts of things the state will find unjust in the foreseeable future. However, to make sense of how the deliverances of a plurality of moral codes could work together

in deciding what injustices a state should remedy, we will need PRIT or at least something fairly similar.

There are three interconnected questions that seem pertinent. The first is the most general: What is the nature of legitimacy? The second: Why does knowing what a state stands for matter? The third: What is the relationship between legitimacy and justifying the basic structure of society? The answer, in different ways, is what I call the *convergence* relation between the goals and methods of a state and the moral codes of its citizens.

I propose that we could define the convergence relation in terms of whether the means and ends characterizing a state (and because there is often no clear distinction between ends and means, hereafter, the states 'meands') can be defended from relevant criticism given the moral codes of its citizens and likewise for the moral codes of citizens regarding the principles that characterize the meands of their state. The intuitive notion here is that we should take a state to be legitimate if its meands coheres with the moral codes of citizens. Identifying the meands of a state presupposes being able to identify the sorts of injustices (and the respective remedies) citizens can reasonably expect the state to so remedy. If the choice to remedy (and how) a sort of injustice can be defended from criticism drawn from the moral codes of citizens, it seems that the state is legitimate for those citizens.

I suspect that this convergence relation, as an account of a suitably thin notion of legitimacy, in fact reinforces a distinction between justifying the basic structure or nature of a liberal state and justifying concrete cases of policy choices. If so, responding to this

216

aspect of Sen's objection, we could find more ways to develop the resources of PRIT to solve larger meta-ethical considerations.

5.1.7 Comparing PRIT with CA

As suggested, with some of PRIT's system sorted out, we could begin to compare how well it stacks up against CA systems in general. I argued in Chapter Two that CA systems generally use a privileged perspective to rank options. If this generally amounts a uniquely correct perspective, foundational pluralism can't be assumed. On the other hand, there is no uniquely correct perspective by which choices in PRIT are ranked. What remains to be seen is whether a CA system could have a plurality of perspectives ranking options.

Appendix

The basic idea behind neighborhood semantics is that a formula, P, is necessary at a world, s, if the truth set of P is one of the sets for the neighborhood function on s. Let me start with perhaps the most intuitive example of how these semantics work. Since the truth set of a conjunction, e.g. $(P \& Q) = (\llbracket P \rrbracket \cap \llbracket Q \rrbracket)$ can be smaller than the truth set of either conjunct, a necessary conjunct at s does not imply either conjunct is also necessary.

Proposition 6.0: $O(P \wedge Q) \rightarrow OP \wedge OQ$ is not valid in the class of neighborhood frames

Proof: Suppose $Sit = \{w, x, y\}$ and that $N(w) = \{\ \{w, x\}, \{y\}\}$. Let $\llbracket P \rrbracket_I = \{y, x\}$ and $\llbracket Q \rrbracket_I = \{y, w\}$. Thus $\llbracket P \rrbracket \cap \llbracket Q \rrbracket = \{y\}$ and so $w \vDash O(P \wedge Q)$. But $\{y, x\} \notin N(w)$ thus $w \nvDash OP$.

Proposition 6.1: $(OP \wedge OQ) \rightarrow O(P \wedge Q)$ is not valid in the class of neighborhood frames

Proof: Suppose $Sit = \{w, x, y\}$ and that $N(w) = \{\ \{w, y\}, \{y, x\}\}$. Let $\llbracket P \rrbracket_I = \{y, x\}$ and $\llbracket Q \rrbracket_I = \{y, w\}$. Again, $\llbracket P \rrbracket \cap \llbracket Q \rrbracket = \{y\}$. So while $w \vDash OP \wedge OQ$ since $\llbracket P \rrbracket_I \in N(w)$ and $\llbracket Q \rrbracket_I \in N(w)$, $\{\ \} \notin N(w)$ and thus $w \nvDash O(P \wedge Q)$.

Proposition 6.2: $O(P \rightarrow Q) \rightarrow (OP \rightarrow OQ)$ is not valid in the class of neighborhood frames

Proof: Suppose $Sit = \{w, x, y, z\}$ and that $N(w) = \{\ \{w, y, z\}, \{w, x\}\}$. Let $\llbracket P \rrbracket_I = \{w, x\}$ and $\llbracket Q \rrbracket_I = \{w, y\}$. Thus $\llbracket \neg P \rrbracket \cup \llbracket Q \rrbracket = \{w, y, z\}$ and so $w \vDash O(P \rightarrow Q)$. But while $\llbracket P \rrbracket_I \in N(w)$, $\llbracket Q \rrbracket_I \notin N(w)$. Therefore $w \nvDash OP \rightarrow OQ$.

Proposition 6.3: $OP \rightarrow P$ is not valid in the class of neighborhood frames

Proof: Suppose $Sit = \{w, x\}$ and that $N(w) = \{\{x\}\}$. Let $\llbracket P \rrbracket_I = \{x\}$. Thus $w \vDash OP$ but $w \nvDash P$.

Proposition 6.4: $(P \rightarrow Q) \rightarrow (OP \rightarrow OQ)$ is not valid in the class of neighborhood frames

Proof: Assume that $w \in \llbracket \neg P \rrbracket \cup \llbracket Q \rrbracket$, so we need to show that $\llbracket \neg P \rrbracket \cup \llbracket Q \rrbracket \notin N(w)$. Let $Sit = \{w, x\}$ and that $N(w) = \{\{x\}\}$. So we need to set $\llbracket P \rrbracket_I = \{\ \}$ and $\llbracket Q \rrbracket_I = \{w\}$, and thus $\llbracket Q \rrbracket_I \notin N(w)$. Therefore $w \nvDash OP \rightarrow OQ$.

Next I show that if sets are defined extensionally, equivalent formula make for equivalent necessities.

Proposition 6.5: $(P \leftrightarrow Q) \rightarrow (OP \leftrightarrow OQ)$ is valid in the class of neighborhood frames with sets defined extensionally.

Proof: Since we assume that $P \leftrightarrow Q$ *iff* $[\![P]\!] = [\![Q]\!]$, if $[\![P]\!]_I \in N(w)$ for any w, then so is $[\![Q]\!]_I \in N(w)$.

I also claimed that if we can aggregate obligations into a obligatory conjunction we can conflate the distinction between inferring permission from obligation and there being no obligations to do the impossible. This claim is classical, i.e., requires explosion and excluded middle, so it's damage is limited, so to speak, in this project. However, one might have noticed that the seriality condition on neighborhood functions suggests a similar issue. Regardless, here is the proof.

Proposition 6.6 There being no obligations to do the impossible classically implies seriality and seriality classically implies there being no obligations to do the impossible.

Proof: Suppose $(\boldsymbol{O}P \wedge \boldsymbol{O}Q) \rightarrow \boldsymbol{O}(P \wedge Q)$ and that seriality fails. Thus, $x \vDash \boldsymbol{O}A$ and $x \vDash \neg \boldsymbol{O}A$ by excluded middle. So, $x \vDash \boldsymbol{O}(A \wedge \neg A)$ by our schema and since $A \wedge \neg A \vDash \bot$ by explosion, $x \vDash \boldsymbol{O} \bot$. To avoid this result, we need the seriality condition (and proving the other direction is trivial.)

Bibliography

G. Bell, D. DeVidi, R. Solomon. *Logical Options*. Peterborough: Broadview Press 2001

K. Binmore, *Natural Justice*. Oxford: Oxford University Press 2005

F. Brandenburg. "Drawing planar graphs on $\frac{8}{9}n^2$ area." *The International Conference on Topological and Geometric Graph Theory* 31: 2008

J. Cardinal, M. Hoffman, V. Kusters. "On Universal Point Sets for Planar Graphs." http://arxiv.org/abs/1209.3594 accessed 10/22/2014

G. A. Cohen. "Incentives, Inequality, and Community" *The Tanner Lectures on Human Values*. Stanford 1991

D. Copp. "Justice and the Difference Principle." *Canadian Journal of Philosophy* 4:2; 1974

M. Dummett. *Truth and Other Enigmas*. Cambridge: Harvard University Press 1980

H. Feigl. "De Principiis Non Est Disputandum (1948)" *Readings in Introductory Philosophical Analysis* ed. John Hospers. New Jersey: Prentice-Hall 1968

J. Forrester, "Gentle Murder, or the Adverbial Samaritan." *Journal of Philosophy,* 1981 :193-196.

N. Fraser. *Justice Interruptus*. Verso 2003

R. Freeman. "Stakeholder Theory of the Modern Corporation"; *Ethical Issues in Business* eds. T. Donaldson, P. Werhane, M. Cording. New Jersey: Prentice Hall 2002

N. Goodman. *Fact, Fiction and Forecast*. Cambridge: Harvard University Press 1983

N. Goodman. *Of Mind and other Matters*. Cambridge: Harvard University Press 1984

C Gowans. "Moral Disagreements." *Moral Disagreements*. D. Brink. New York: Routledge 2000

H. Grice. "Logic and conversation." *Syntax and Semantics* 3 eds. P. Cole and J. Morgan; 1975

J. Harsanyi. "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory." *The American Political Science Review* 69:2; 1975

R.M. Hare. *The Language of Morals*. Oxford University Press 1991

A. Hasse. "Keeping Wolff from the Door: California's Diminished Capacity Concept." *California Law Review* 60:6; 1972

R. Hilpinen. *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics* Volume 152. "Non-Kripkean Deontic Logic". P. Scotch, R. Jennings. E. Berlin: Springer 1981

---- "The Intension of Human Rights." *Atlantic International Studies* 5; 2006

D. Keyt. "The Social Contract as an Analytic, Justificatory and Polemic Device." *Canadian Journal of Philosophy* 4:2; 1974

D. Lewis. *Counterfactuals.* Cambridge: Harvard University Press 1973

R. Lippke. "Towards a theory of prisoners' rights." *Ratio Juris* 15; 2002

E. Mares. *Relevant Logic.* Cambridge: Cambridge University Press 2004

T. McConnell. "Moral Dilemmas and Consistency in Ethics"; *Canadian Journal of Philosophy*, eds. B. Chellas, J. King-Farlow, K. Nielsen et al. Volume VII, Number 2. Canadian Association for Publication in Philosophy, 1978

T. McConnell "Moral Dilemmas and Requiring the Impossible." *Philosophical Studies 29(5)* June 1976 pgs. 409-413

R. Nozick. *Anarchy, State, and Utopia*. Basic Books 1974

D. Nute. "Defeasible Reasoning and Moral Dilemmas" *Defeasible Deontic Logic* Volume 263; eds. R. Koons, T.K. Seung, Boston: Kluwer Academic Publishers 1997

B. Orend. *Human Rights*. New York: Broadview Press 2002

P. Pettit. *Politics, Philosophy, Economics* Volume 4 pg 157 "Rawls's political ontology" 2005

G. Priest. *Doubt Truth to be a Liar*. Oxford: Oxford University Press 2006

G. Priest. "Logical pluralism hollandaise." *The Australasian Journal of Logic* 6; 2008

G. Priest. *An Introduction to Non-Classical Logic*. 2nd Edn. Cambridge: Cambridge University Press 2008b

J. Rawls. *Political Liberalism*, extended edn. New York: Columbia University Press 1996

J. Rawls. *A Theory of Justice*. Cambridge: Belnap Press of Harvard University Press 1999 [1971]

G. Restall *An Introduction to Substructural Logics*. New York: Routledge 2000

T. Scanlon. *What We Owe Each Other.* Cambridge: Harvard University Press 1998

T. Scanlon. *Moral Dimensions*. Cambridge: Harvard University Press 2008

A. Sen. *The Idea of Justice*. Cambridge: Harvard University Press 2009

P. Singer. *Practical Ethics*. Cambridge: Cambridge University Press 2011

C. Tappolet. " Mixed inferences: a problem for pluralism about truth predicates."
Analysis 57; 1997

C. Tappolet. "Truth Pluralism and Many-Valued Logics: A Reply to Beall." *The Philosophical Quarterly* 50; 2000

A. Tarski. "The Concept of Truth in Formalized Languages." *Logic, Semantics, and Metamathematics*; 1956

J. van Benthem, D. Grossi, F. Liu *On the Two Faces of Deontics: Semantic Betterness and Syntactic Priority*. http://www.illc.uva.nl/Research/Reports/PP-2011-29.text.pdf
Accessed Aug 7, 2013

B. Van Fraassen. *The Scientific Image*. Oxford: Oxford University Press 1980

T. Ward, A. Birgden. "Human rights and correctional clinical practice." *Aggression and Violent Behavior* 12; 2007
J.R. Welch Rodopi "Two Types of Moral Dilemma" *The Future of Value Inquiry* Volume 112. eds. M. Hayry, T. Takala: BV Editions 2001

B. Williams. *Moral Luck*. Cambridge: Cambridge University Press 1981

B. Williams. *Morality*. Cambridge: Cambridge University Press 2012