

5-2018

# Damned If You Do, Damned If You Don't: A Logical Analysis of Moral Dilemmas

Samuel Monkman  
smonkman@bowdoin.edu

Follow this and additional works at: <https://digitalcommons.bowdoin.edu/honorsprojects>



Part of the [Ethics and Political Philosophy Commons](#)

---

## Recommended Citation

Monkman, Samuel, "Damned If You Do, Damned If You Don't: A Logical Analysis of Moral Dilemmas" (2018). *Honors Projects*. 98.  
<https://digitalcommons.bowdoin.edu/honorsprojects/98>

This Open Access Thesis is brought to you for free and open access by the Student Scholarship and Creative Work at Bowdoin Digital Commons. It has been accepted for inclusion in Honors Projects by an authorized administrator of Bowdoin Digital Commons. For more information, please contact [mmcderm2@bowdoin.edu](mailto:mmcderm2@bowdoin.edu).

Damned If You Do, Damned If You Don't:  
A Logical Analysis of Moral Dilemmas

An Honors Project for the Department of Philosophy  
By Samuel J. Monkman

Bowdoin College, 2018

©2018 Samuel J. Monkman

# Table of Contents

Acknowledgements	p. ii
Chapter 1 - Introduction to Moral Dilemmas and Consistency	p. 1
Chapter 2 - Introduction to Deontic Logic	p. 19
Chapter 3 - Agglomeration and Ought-Implies-Can	p.33
Chapter 4 - Deontic Consistency and Entailment	p. 52
Chapter 5 - Conclusion and Implications	p. 66
References	p. 87

## Acknowledgments

I would like to thank Bowdoin's Philosophy department for supporting my interest in this project and academic philosophy in general. It has truly been a joy to work with such thoughtful scholars and dedicated educators over the past years. In particular, Professors Sarah Conly and Scott Sehon have provided endless guidance throughout this process. This project would not have been possible without their supportive and critical approach. I have learned a great deal over the course of the project, and the great majority of that learning is due to their attentive advising.

I would also like to acknowledge the support I have received from my fellow students. Many students pursue honors each year, and I have found a unique sense of togetherness in this community. Their shared passion and commitment to academic rigor and excellence has inspired me countless times over the past year. I have been truly humbled to work alongside such intelligent and dedicated people.

## Chapter 1 - Introduction to Moral Dilemmas and Consistency

The term “moral dilemma” has a number of uses in common and philosophical speech. This project is not concerned with the common use of “moral dilemma” to describe situations where it is difficult to determine what is the right thing to do, or where morality conflicts with pragmatic concerns. Those situations are interesting in terms of their impact on applications of ethics, but I am more concerned here with dilemmas of a different sort. A moral dilemma, as I use the term, is a situation in which an agent has an obligation to do some action A and an obligation to do some action B but cannot perform both actions. In particular, I am interested in answering the following question: Is it possible for a consistent moral theory to produce genuine dilemmas?

I will clarify just what I mean by “genuine dilemmas.” It should be uncontroversial to say that many moral theories give *prima facie* dilemmas, situations where an agent has two *prima facie* obligations that stand in conflict to one another. These situations are typically resolved with a sort of moral calculus. One obligation is often seen as stronger or more urgent than the other, and so the morally correct course of action is to satisfy the superseding duty. A genuine dilemma, then, is an all-things-considered dilemma. In other words, a situation qualifies as a genuine moral dilemma when it involves a conflict between two all-things-considered obligations. In these cases, neither obligation can be discharged by appeal to the other or by any external factors.

All-things-considered dilemmas are called “genuine” because they present a more serious challenge to a moral theory than do *prima facie* dilemmas. It is to be expected that any theory that posits certain general obligations, like the obligation not to kill or the obligation to keep one’s promises, will admit some situations where two or more of these obligations are in apparent conflict. For example, a soldier who is ordered to attack and kill his enemy faces a *prima facie* dilemma under any theory that posits a duty not to kill and a duty to follow orders from legitimate authority. Most plausible theories will include a system for analyzing such cases, and many will involve a preference in priority between the two duties. More pacifist theories will tell the soldier it is permissible to disobey his orders when those orders demand he commit wrong (especially violent) acts. Some other theories will value honoring contracts and authority more highly, and allow that the soldier suspend his duty not to kill when the killing is justified (in this case by obedience to legitimate authority).

There are some theories, however, that would say of the above case that neither action is right. The soldier has an obvious responsibility not to kill other people. He also has a genuine duty to follow orders, especially since he has entered into a contract with his government to do so in all circumstances. If a theory does not include a method for deciding between conflicting obligations in every possible case, it will admit some genuine dilemmas. In the soldier’s case, he will do wrong no matter which choice he makes.

I will now turn to the other most important feature of this project: Consistency in ethical theories. Theories that admit genuine dilemmas strike many

as logically inconsistent. Kant said in *The Metaphysics of Morals* that “a conflict of duties and obligations is inconceivable” (17). Of course, Kant has had direct influence on many moral theorists since, and a large number continue to echo this sentiment. The idea that a consistent basis for generating ethical duties can give rise to conflicting obligations seems strange. Several contemporary thinkers have defended dilemma-producing theories, however, including Ruth Barcan Marcus and Bernard Williams. For both of them and many others, the argument hinges on consistency. The question remains, then, of what exactly one means when one says that a theory is consistent or inconsistent.

When one creates a moral theory, it has a particular structure. The theory begins with certain moral principles, a “moral code” as Marcus calls it (123). These principles give general duties and permissions. They require generic agents to fulfill certain obligations at whatever time the situations arise. Principles may also be conditional, in which case they apply only to those agents who meet certain criteria. For example, parents may be morally required to provide for their children where others are not. These principles are something like mathematical equations with unassigned variables, and correct moral action will satisfy the principles for every relevant assignment of agent, time, place, and context for action.

Theories consist (at least in part) of a moral code and its rules of application. For example, a utilitarian theory has just one fundamental principle: Agents must act to maximize utility (which will be defined differently under different versions of the theory). The rules of application are that in any situation, an agent should prefer a course of action with greater resultant utility. In cases where two actions have

equal utility, the choice is a matter of indifference. Other theories may have somewhat more complicated formulations, but this structure is close to universal for theories of morality that hinge on action. Theories of moral character and virtue will differ somewhat, as they are often not prescriptive theories *per se*.

Given this structure for moral theories, the reader may see how a description of consistency follows. An agent considers the correct application of the moral code, and rules a theory inconsistent if this application results in a logical contradiction (in a properly defined logical system). This is a weak test for consistency of any sort of theory. Arguments for inconsistency are done *ad absurdum*. The skeptic assumes the theory is correct and shows that its application yields an impossible result. This shows that the theory's foundation is faulty and must be revised or abandoned. Less formal interpretations of "consistent" apply in some situations, but I will restrict my use to this more rigorous sense of the term.

How does this impact my talk of dilemmas? The opponent of theories that allow dilemmas will wish to dismiss them by this process. The argument will assume that a generic dilemma occurs, and show by application of certain incontrovertible principles that a contradiction follows. When I use phrases like "opponent of dilemmas" or "defender of dilemmas" throughout this project, I mean those who fall on one or the other side of the debate about the consistency of dilemma-admitting theories. In order to argue for either side of this issue, one must have a system for formalizing ethical claims. In the next chapter, I will discuss that endeavor in depth. Following that, I will give the two most compelling arguments attempting to show dilemma-producing theories to be inconsistent. Both of these



arguments will rely on certain axioms of normative logic, representing principles fundamental to *any* plausible moral theory. I will argue in some cases that the principles are not in fact necessary to represent a normative reasoning. In the other case, I will show that the axiom in question does not apply to the relevant situations. Before I dive into the problem of symbolization and formal argument, however, I will discuss in greater depth the history of moral dilemmas and certain efforts to show by conceptual argument whether they are possible.

One of the earliest often-cited examples of a moral dilemma comes in Plato. In the *Republic*, he proposes a response to a possible definition of right action. “Truthfulness and giving back anything one has borrowed from someone,” according to Plato, is an inadequate characterization (331c). Socrates, Plato’s mouthpiece in the text, considers the case where he borrows a weapon from a friend. Before he returns the weapon, he learns that his friend has gone insane, and so will likely do significant harm with the returned weapon (Plato 331c). In this situation, Socrates faces something of a dilemma. He can honor his promise, as Simonides’ stipulated definition of “justice” would require. If he does so, however, he will be indirectly responsible for his friend’s wrong actions. He will certainly be culpable if we believe he has an obligation to prevent whatever harm his friend will do. On the other hand, if he does not return the weapon, he will clearly be breaking a promise, which Simonides and most modern thinkers accept is at least *prima facie* wrong. Plato believes this situation, at least, is not an example of a genuine dilemma. Rather, he believes it is fairly plain that Socrates must break his promise, and so the

case refutes Simonides' conception of justice. I think most would be inclined to agree.

Socrates' case represents the fairly common practice of referring to any situation of a complex or difficult moral nature as a dilemma. Under some theories, Socrates would in fact be facing a genuine dilemma, but many thinkers prior to the mid-twentieth century maintained that any such situation would admit of a right choice. That is, any apparently dilemmatic case can be analyzed away when one discovers which of the duties took precedence. These situations are often quite dramatic, earning the title "tragic dilemmas." One need not look far in classical literature for such a case. Perhaps the most popular example is that of Agamemnon's choice to sacrifice his daughter Iphigenia. He is presented with the following dilemma: He intends to sail his fleet, but the winds are dead. He discovers that the goddess Artemis has stopped them from blowing because he has slighted her. He can only restore the winds if he sacrifices Iphigenia. If he does not do so, he will betray his promise to aid Menelaus in recapturing Helen (Euripides).

Agamemnon's case is certainly tragic in every sense of the word. He is forced to choose between his honor and his daughter, and in the end he does indeed choose to sacrifice Iphigenia. Oddly, to the modern reader this seems clearly to be the wrong choice. One would hope that Agamemnon valued his daughter's life more highly than his promise to wage war over Menelaus' pride. Again, this case seems less to reflect a dilemma in moral terms, but rather to showcase how moral duties often conflict with other important considerations, such as personal honor. These sorts of situations are different from genuine moral dilemmas, but their use in

classical examples of morality is important. A preoccupation with apparent dilemmas of the sort that appear in Plato's and Euripides' works may have had some effect on the frequent dismissal of dilemmas as an important concern for moral theory. The most common examples of dilemmas are not real examples, in that they do not seem to pose a choice between two equally weighty obligations. In these situations, two *prima facie* duties conflict, but there is little ground in either case to argue that neither obligation can be overridden.

Despite their apparent moral simplicity, these classical cases are noticeably at odds with a particular sort of morality that is actually quite popular in modern ethical theory. Kantian ethics is founded upon the notion that a single moral rule (the Categorical Imperative) can be applied to determine the moral status of any considered action. The unity of this rule is paramount. According to Kant, the Imperative, which is given in terms of several formulations, is just one rule presented in various equivalent forms (Cureton and Johnson 9). Since Kant believes moral truths to be matter of necessity, he rejects the notion that any two duties can conflict, even if one is more pressing and overrides the other (Timmerman 40-41). A necessary truth cannot be shown false by appeal to other facts or truths, as would be the case if one allowed certain moral truths to override others (according to Kant). Under a strict Kantian view, even *prima facie* dilemmas may be inadmissible, since the application of the Categorical Imperative must not yield mixed results. A correct application of the Imperative should always perfectly guide action. That is to say, the strict Kantian will insist that in every situation a course will be available which satisfies the Categorical Imperative.

What might the Kantian say about Socrates' and Agamemnon's cases? Most likely, she would argue that in neither case is there a conflict even of general principles. Since Kant's Categorical Imperative is a unified founding principle for his moral theory, it seems intuitive that it should never generate conflicts of duties. One may be inclined to say that a consistent ethical principle should not yield conflicting results. In particular, if one accepts Kant's claim that a conflict of morality is rationally impossible, it should be straightforward to say that any apparent dilemma will be a result of mistaken reasoning on the agent's part (when they reason using the Categorical Imperative). One may internalize certain general guidelines for right action, but upon carefully considering a situation, the Imperative should always leave at least one acceptable course of action. Since a single principle provides a complete moral guide, correct application of the principle will always give correct moral answers. Taken together with Kant's insistence that duties cannot conflict, this means that the Categorical Imperative must never return a dilemmatic result. If one accepts this, then neither Socrates nor Agamemnon faces a conflict even of general moral principles. Kantian ethics gives only one ultimate principle, so one cannot say that two or more principles conflict with one taking precedence. Rather, the apparent conflict is only with the agent's mistaken application of the Imperative.

Of course, Kantian theory must make room for the possibility of complex and difficult moral situations. Mark Timmons maintains that instances of *prima facie* conflict are in fact possible on the right interpretation of Kant. He presents the following situation:

*Lizbeth's Lie:* Lizbeth is hiding an acquaintance, Tom, who is wrongly believed by some murderous gang members to have ratted out one of the gang to the police. A gang member shows up at Lizbeth's front door asking her for the whereabouts of Tom. To tell the truth would surely lead to the murder of Tom, and so would (let us suppose) just refusing to answer. So Lizbeth lies, saying that she doesn't know where Tom is (Timmons 218).

One of the most straightforward duties generated by the Categorical Imperative is the obligation not to lie. The first formulation of the Imperative gives a test for an action's moral status: If the agent can will their action as a universal maxim in a rational universe and could rationally take that same action in that universe, the action is permissible (Cureton and Johnson 5). If one wills a universe in which lying is a universal law (that is, everyone lies all the time), it would be irrational to lie about Tom's whereabouts, since the gangster would never believe the lie. Thus, Lizbeth has a duty not to lie. On the other hand, Kantian ethics includes a duty to promote others' happiness (Timmons 214). Protecting Tom's life seems quite a weak result of protecting his happiness, so Lizbeth also has a duty not to reveal his location and facilitate his murder. The potential conflict of duties is obvious. It is intuitively obvious that Lizbeth acts rightly when she lies to protect Tom, so in this case one duty takes precedence. Timmons gives what he thinks would be Kant's reasoning for the choice. A duty does not override another based on the strength of the obligations, but rather based on the strength of the grounds generating the obligations (Timmons 218). In this case, there are grounds for an obligation not to lie and grounds for an obligation to rescue Tom. Timmons

maintains that the grounds for the obligation to save Tom are stronger, and therefore one obligation obtains, while the other fails to apply.

It seems Kantian theory must have some way to explain the apparent conflict of such fundamental principles as those at play in Timmons' example. Timmons' own proposed reading of Kant seems quite reasonable, and it allows the Kantian to grant what seems so obvious intuitively: There are situations in which an agent will have two moral duties in conflict, although one or the other of the duties may be overridden. If this is the case, what becomes of the notion that a single unified principle must not give conflicting results? The answer to this will have implications for other moral theories, such as utilitarianism. Maximizing utility is another singular moral principle, and so it is subject to much the same intuition at play in the case of Kantian ethics.

It seems now that this intuition runs into some problems, however. It will be helpful to consider contemporary theories with more than one founding principle. In pluralist moral theories, "a plurality of equally basic morally relevant features" must be considered when evaluating an action's moral status (Timmons 246). These theories differ in some important ways from monistic theories, like Kantian or utilitarian ethics. Nonetheless, they are subject to many of the same general considerations.

Just what is a pluralistic theory? I am concerned specifically with what is called "foundational pluralism" (Mason 1.1). This is the view that a correct moral theory will not be founded upon a single unifying theory of value. Rather, actions and outcomes are assigned value in reference to a number of metrics. For example,

Judith Jarvis Thomson argues that all judgments of goodness (i.e. of value) really evaluate “goodness in a way,” that is to say by reference to some particular paradigm of value (276). No objective property of goodness exists on this view. Where Kant’s and similar views appeal to a singular unified notion of value, a pluralistic view will maintain that no such unity holds for genuine moral reasoning. Rather, moral agents do and should appeal to different sources or types of value in their different judgments of moral value and status.

The most pertinent feature of moral theories to my present project is consistency. A theory of any sort will be unsatisfactory if it is found to be logically inconsistent, and inconsistency is the most common charge against moral theories that admit dilemmas. The worry arising from Kant’s discussion of conflict is related to consistency. If a single foundational principle (like the Categorical Imperative) produces dilemmas, it at least gives the impression of inconsistency. It is not clear that such a situation actually implies a logical contradiction, however. On that strict use of “consistency,” the status of dilemmas remains to be seen. As I will argue in later chapters, dilemmas’ existence only generates a logical contradiction if certain questionable axioms are adopted. These axioms will not hold in some plausible moral theories, so they must not be taken as fundamental to an ethical logic. In that case, producing dilemmas should not be taken as a sign of a theory’s inconsistency, unless the theory’s own conceptual principles would suggest that dilemmas are inadmissible. Monistic theories may or may not have this last feature, but the intuition remains that a unified moral principle should guide action perfectly. By

examining pluralistic theories, one can discover why it might be that even a monistic theory might admit dilemmas without rendering itself inconsistent.

Do we have the same anti-dilemmatic intuitions about pluralistic theories?

There are two ways to approach a pluralistic moral theory. On the one hand, one could view the various foundational principles as completely separate from one another for all theoretic purposes. For example, if I have a duty to care for my family, that duty may be entirely separate from my duty to lessen the suffering of distant strangers. I could very well have both obligations, and they would presumably be at odds in some situations. Another way to approach pluralistic moral theories would be to say that such a theory resembles a monistic theory in many ways. The agent might then treat the set of foundational principles as a single conjunctive principle. Instead of a collection of separate principles or systems of value, the theory's basis is represented as a number of ways in which the single property of goodness manifests itself. The Categorical Imperative itself may be read this way. If one does not take Kant quite at face value when he claims that the formulations of the Imperative are strictly identical, it is plausible to claim that each formulation of the Imperative is in fact a different fundamental principle. Together they capture a unified system of value, which merely shows itself in several different action-guiding rules. It is taken as a single unified moral law, which will give different guidance in different situations based on its constituent sub-principles. Each of these two interpretations of a pluralist theory will involve different attitudes toward dilemmas.



The first, more genuinely pluralistic view is often friendly to conflicts of duty, even to genuine dilemmas. A popular example comes from Jean-Paul Sartre. A student's brother is killed in the German offensive (in France) in 1940. The student wishes to avenge his brother's death and fight the Nazi forces, whom he rightly considers evil. However, he lives with his aging mother, and he is her only joy in life (McConnell 1). If we accept the implicit stipulation that both of these desires have moral motivation, a dilemma arises. The student has a duty to care for his mother, but he may also have a duty to fight against the oppressive and evil Nazis. Under this description, it seems he has two different types of obligation. He has a familial obligation, with "limited scope but certain efficacy" to provide comfort to his mother (McConnell 1). He also has perhaps a nobler but also more ambitious duty to join a fight that may save millions of lives from persecution. If the student accepts a moral theory that closely resembles common sense morality, he will likely accept both of these duties. Given Sartre's use of the case, it seems these are completely distinct principles of morality. In that case, at least a *prima facie* dilemma obtains. It may be possible to assign a weight to the conflicting principles, in which case a genuine dilemma would not arise. Rather, the student would follow the obligation whose motivating principle was stronger. It is not clear that either the duty to fight tyrants and murderers or the duty to care for one's family will take precedence in every case, however, so it is difficult to say one of the principles lends its generated duties greater weight than the other. If that were the case, a genuine dilemma would arise. Whatever the student does, he will fail to follow one of the other genuine moral principle.

When one considers an apparent dilemma in a conjunctively-founded moral theory, the situation more closely resembles Kant's considerations. Similarly to Lizbeth's case, Sartre's student applies his moral principle to determine what to do in this situation. Instead of testing his choices against a single rule, he checks each possible action against each conjunct in his conjunctive theory. As in Lizbeth's case, he discovers that both options open to him are required by at least one conjunct. Thus, he faces a potential dilemma. As in the above case, he may resolve the situation by weighing his sub-principles against one another and deciding which is more pressing or more powerful. This approach may run into the same problems I mentioned in the previous paragraph, since the principles in question still do not present an obvious choice. In any event, the approach of treating a pluralist theory as a conjunctive monistic theory presents the same apparent problem of consistency.

It seems intuitive to say that a theory with multiple founding principles is really based on all those principles taken together. If there is to be any test for inconsistency in the theory as a whole, it will certainly consider all the principles together. That is to say, a pluralistic theory will be treated as conjunctive if it is to be taken as a single theory at all. Even in a theory like Thomson's, which is based on a plurality of types of value, one might focus on those actions and outcome which have some moral component, treating the various reference frames of goodness as subtypes of moral value. In any case where one cannot make this move, it will be reasonable to say that the object in question is not really a single theory but multiple theories of morality, and separate theories need not be mutually consistent.

It seems that in Sartre's case, a genuine conflict of principles arises. The conflict may be ultimately *prima facie*, it is much harder to see the right course than in most of my previous examples of *prima facie* dilemmas. What does this mean for the consistency of moral theories on the whole? I have argued that any pluralistic theory will be conjunctive if it is to be a single theory at all. In that case, pluralistic theories resemble monistic theories in several ways, especially in terms of testing consistency. As such, it would be reasonable to expect that one sort of theory will admit dilemmas if the other does, at least if we take "admits dilemmas" to mean "does not imply a contradiction when dilemmas are posited." It is clear that pluralistic theories admit *prima facie* dilemmas, so monistic theories may do the same without danger of rendering themselves inconsistent. In the end, the question comes down whether in situations of apparent conflict one duty can always be given preference over the other.

Kantian theory and many pluralistic theories share a certain form. They mostly fall into the category of deontological ethical theories. Typically, these theories involve mandates about types of action (such as lying or saving a life), rather than giving a principle that applies straightforwardly to every specific action. Utilitarianism takes the latter route. This differentiates it from other theories. The principle of maximizing utility (or good consequences) is not a restriction on certain types of action. Rule utilitarianism will more closely resemble a pluralistic theory in this regard, since it selects those guiding principles that give the best utilitarian results. The only situation resembling a dilemma that will arise for a strict act utilitarian is one in which no unique maximum of good consequences exists.

Assuming an agent has finitely many courses open to them at any given decision point, this means that two or more actions share the same utility value. In that case, consequentialists are generally happy to leave the choice between the maximizing actions up to non-moral factors. As long as the agent takes one of the courses with the most positive result, the choice between those actions is a matter of indifference. Thus, dilemmas do not arise, because there will never be a situation in which an agent must take action A and must take action B, but cannot take both in conjunction. The theory is designed to accommodate such cases by saying that when the principle of maximal utility generates two incompatible obligations, the agent is only required to take one or the other of the actions, and there is no moral reason to choose one over the other.

The theories I have discussed thus far are examples of what Ruth Barcan Marcus calls “ethical formalism” (124). She differentiates these theories from ethical intuitionism. Formalism maintains that a sound moral theory consists of a set of principles that apply to specific cases. Morally analyzing a decision point involves testing the available actions against the principle or principles underlying the moral theory. An intuitionist theory, on the other hand, will maintain that no set of principles and rules of application can adequately capture moral reasoning (Marcus 124). The moral intuitionist denies that dilemmas can be solved by appeal to a strict ordering of fundamental principles, since he will deny that any such ordering is possible. Rational intuition about specific cases is the final arbiter of moral status, while theoretic principles provide *prima facie* moral reasons.

From this it follows that when application of principles gives a dilemma, it is only a *prima facie* dilemma. According to Marcus, the intuitionist will maintain that such a situation can be resolved by recognizing the conflict is “epistemological and not ontological” (124). The apparent dilemma is a matter of uncertainty, rather than a genuine conflict of obligations. Intuitionist theories need not be concerned with consistency in the same way formalist theories are. The intuitionist does not accept that morality arises out of the application of fundamental principles, so there is nothing in an intuitionist theory to call inconsistent. One may reason wrongly about cases, but this is often a case of incomplete knowledge, or of taking heuristic principles too seriously. Intuitionist and formalist theorists differ in their approach to dismissing dilemmas, but historically most theories have been unwilling to accept genuine dilemmas into their moral ontology.

A number of moral theories do not admit dilemmas on the grounds that doing so would render the theory inconsistent. This stems largely from an intuition that consistent principles will not give rise to conflicts. However, most plausible theories present agents somewhat regularly with *prima facie* dilemmas. In other words, consistent bases for morality frequently generate conflicts of duty when every duty is taken to be equally pressing. In that case, what remains is to determine whether every theory can be organized such that its founding principles are ordered according to moral strength, or perhaps to elaborate on basic principles to account for potential dilemmas. Another option would be to adopt an intuitionist theory. Either of these approaches may dismiss the problem of dilemmas, but it is not clear for many theories whether their founding principles or general duties can be

plausibly ordered, and intuitionism runs into a number of problems, which I will not discuss at length here. It is enough that it now seems plausible that dilemmas may not be as problematic as they first seem. I will now proceed to give a general account of the formalization of ethical claims, called “deontic logic,” and then to discuss the two most widely-discussed and most compelling formal arguments for the inconsistency of dilemma-producing theories. In both cases, I will argue against the principles used in the argument, or at least in their relevance to dilemmas.

## Chapter 2 - Introduction to Deontic Logic

I will begin here with a discussion of deontic logic generally. I am aiming for the reader's grasping the basic concepts used in the few proofs I will give, so this will be a very simplified account of a complicated subject, with an eye toward understanding the semantic and interpretive elements rather than the more technical aspects. "Deontic logic" refers to formal symbolic systems used to represent normative statements. The field of deontic logic has developed largely as an adaptation of modal logical systems. As the name suggests, modal logic attempts to formalize and understand claims of possibility and necessity. Deontic logicians developed parallel systems to capture normative reasoning, especially in arguments about ethics. It has important applications in understanding the structure and relationships of ethical obligations. Deontic logic is of particular interest to the project at hand, since I am concerned with the logical status of dilemmas across all plausible moral theories. The correct approach to this problem is to consider which formal axioms must be admitted into any normative logic and determine the status of dilemmas under the axioms.

The first question to raise is this: Why do we need deontic logic in the first place? How does it improve our understanding and reasoning about ethics (and normativity more broadly)? The primary motivation matches that behind any logical system. By developing formal rules for reasoning about norms, one is better able to reach correct inferences in normative arguments. A formally defined symbolic system will allow statements of ethical duties and permissions to be

paraphrased into schemata, which are subject to well-defined rules of inference. In a useful system, nearly every statement in ethics will have a paraphrase. Thus, argument can be handled formally, with a clear decision process for determining validity. It should also be noted that such a system must capture normative claims in a way first-order logic cannot. Imperative statements are not propositions, of course, and so it is not possible in many cases to accurately paraphrase an ethical argument relying on the notion of moral imperative into first-order logic. A semantic system will be required which reflects normative force.

In particular, the project of developing deontic logic is the project of capturing implication, mutual incompatibility, and similar notions as they pertain to imperatives. Deontic logic is used almost exclusively as a logic of ethics, but in principle it should apply to any normative system. In this way, the fundamental principles of deontic logic are prior to ethical theory. Deontic logic captures imperatives and normative value as a whole, and those statements that are logically valid in the system will be those that are true because of the nature of normative statements. No particular obligations or permissions will be derivable as theorems in the logic, because these rely on specific conceptual parts of ethics. Rather, any ethical theory will introduce certain additional principles to the fundamental ones. These basic axioms of a deontic system represent principles of normativity without which a set of norms would be unintelligible or deeply inconsistent.

A formal deontic system will help with theoretical reasoning about the nature of ethics and norms, but it also enables deeper analysis of applied ethics. A clear formal logic should help to clear up problems involving counter-to-duty or



secondary obligations Consider the following case: You pass by a pond in which a man is drowning. Due to your fear of water, you cannot bring yourself to jump in and save the man. In this situation, you have presumably violated a moral duty. If one is able to save another's life at minimal risk to oneself, one should do so. It is clear that you have an obligation to call for help, given that you will not jump in to save the man yourself. Certain logical questions arise when considering counter-to-duty imperatives such as this, however. In this case, our intuition is clear that one must do whatever one can (perhaps without seriously endangering oneself) to rescue the drowning man, hence the obligation to dive into the pond. Presumably, the non-ideal agent is still obligated to save the man, but they are unwilling. In that case, another obligation arises, now to ensure his rescue by someone willing and able. The intuitive solution is obvious, but in formal argument it does well to be able to challenge and confirm intuitions. A carefully designed deontic logic will accommodate analysis and explanation of obligations that arise in non-ideal situations. Indeed, this is one of the central issues of deontic modality. I will discuss this particular semantics of deontic logic shortly.

The above example leads me to a more theoretical reason for introducing deontic logic. The methods involved in creating and applying such a logical system will give significant insight into the structure and nature of specific obligations. In the drowning-man example, the obligation to call for help is secondary to the obligation to save the man. Logical analysis illuminates the relationships between interdependent and counter-to-duty obligations. A logic specific to ethics will avoid the problems with conditionals that result from relying strictly on first-order logic.

This use should extend beyond counter-to-duty situations, and address several apparent puzzles in intuitive moral reasoning.

Consider a case in which a person is voting to elect a mayor for their city. One candidate has promised to allocate funds for the improvement of education and healthcare. This candidate is known to be trustworthy. The other candidate has made some vague promises about cutting taxes and supporting business, but has a dubious record of fulfilling promises and has presented no clear plan. The first candidate would clearly be preferable, but whether the matter comes down to a moral obligation is not obvious at first glance. It would be in line with intuition to say that voting for a better candidate is morally required, but from where exactly does that obligation arise? In this case, it will not be out of violation of some primary duty, but rather a specific application of a more general duty. Perhaps we believe that we have an obligation to maximize our fellow citizens' well-being, or at least to ensure that our and others' tax dollars are not misappropriated by a crooked mayor. In either case, a general obligation to pursue some good or prevent some evil may generate a duty to make a certain mark on a certain piece of paper. To determine whether such a duty arises, one would consider the general obligation in question and apply rules of inference to determine whether, given the facts about expected outcomes and applicability of moral requirements to specific types of action (in this case, voting), and then decide whether or not one has a genuine ethical obligation to vote for one candidate or the other.

By generalizing this process, a theorist may understand a little better the mechanisms by which duties give rise to one another. In an uncertain case, the

argument for a certain obligation can be given in formal terms and examined for validity and soundness. One of the most powerful tools in this situation will be argument by contradiction. A common argumentative tactic is to assume one's opponent is correct, or that a certain action is required, and then demonstrate that the situation would lead to a logically impossible result. For this to work in the moral case, a thinker must understand just what constitutes a normative contradiction. It will not be as straightforward as " $p \ \& \ \sim p$ ", as one might hope. A more robust logical system will be required. Once a theorist has the tools to apply strict, decidable rules to arguments, he or she will be able to give more certain accounts of the specific obligations of individual actors. In short, a system of deontic logic will dispel a great deal of faulty reasoning from ethical discourse, since it will both require and increase clarity and agreement at least in methods of inference. A keener understanding of the logical structure of moral claims may also engender a more structured approach to their metaphysical status. As I will discuss in greater detail as I consider specific systems for deontic logic, the axioms and theorems we admit into a system will depend on and inform how we think about the nature of norms, and especially how norms relate to one another and to our own actions.

With these motivations in mind, I will now turn to the workings of deontic logic itself. Deontic logic is closely related to modal logic in both its history and construction. Modal logic deals with claims of possibility, necessity, and contingency, called "alethic modality." Where first-order logic is able to capture reasoning only about real things and events, modal systems can paraphrase arguments from counterfactual claims. It also allows one to distinguish between

necessary truths and those that hold only contingently. For example, in first-order logic, the paraphrase for, "Every bachelor is unmarried" will not be logically valid. Using "Bx" for "x is a bachelor" and "Mx" for "x is married," the paraphrase looks like this: " $(\forall x)(Bx \supset \sim Mx)$ ." This schema will require premises to be proven true, and it is no logically different than the statement "All ravens are black." That is to say, the notation and syntax do not indicate any difference between logically true and contingently true statements. To solve this problem, we introduce new operators " $\Box A$ " for "A holds necessarily" and " $\Diamond A$ " for "A is possible" (Girle 3). Then, we can say " $\Box(\forall x)(Bx \supset \sim Mx)$ ," which means "All bachelors are necessarily unmarried." The modal operator allows one to indicate that this proposition holds under all possible contingent states of affairs.

I will clarify this last statement somewhat and briefly describe the formal semantics behind modal logic. The standard model is called "Kripke semantics," though in actuality it was developed in parallel by a number of logicians in the mid-twentieth century (Wolenski 273). The principle is that the operators quantify over possible worlds, or universal states of affairs. The necessity operator is interpreted as "in all possible worlds." Thus, any state of affairs that holds necessarily is said to hold across all possible worlds. Similarly, the possibility operator says "in some possible world," so that " $\Diamond A$ " means the same as " $\sim \Box \sim A$ ." "A" holds in some possible world if and only if " $\sim A$ " does not hold in every possible world. There is an important notion in the development and use of modal logics called a "world access relation." This relation allows one to adapt these logical systems to various types of modality. One might be concerned with causal, rather than logical possibility, for

example. I will not delve too deeply into the specifics of this relation and its properties. Though it is important to a nuanced grasp of modal logic, I am here more concerned in giving the reader a general idea of how these systems are constructed.

Now, with a basic understanding of modality and formal systems, I shall proceed to discuss the topic at issue for this project. Deontic logic, as I said before, is logic about norms. It is used specifically for ethical reasoning, but the systems involved can easily apply to other normative models. A game with set rules might make use of a primitive deontic logic, for example, and since we often treat legal and civil duties similarly to moral ones, we will find that the same rules of reasoning often apply. The standard approach to deontic logic is essentially deontic modality, though it differs from other modal reasoning in several ways. The operators “ $\Box$ ” and “ $\Diamond$ ” are replaced by “ $O$ ” and “ $P$ ”, with “ $O$ ” standing for obligation and “ $P$ ” for permission. This difference is mostly cosmetic, but it allows arguments in deontic logic to make simultaneous use of alethic modality without undue confusion of symbols.

What I mean by the phrase “deontic modality” needs a bit of explanation. If modal logic is the logic of possibility and necessity, then deontic logic concerns deontic possibility and necessity. Under the ordinary “possible worlds” interpretation, this means that each permissible action generates a deontically possible world. In each deontically possible world, every obligation can be fulfilled jointly. This interpretation mirrors the way possible world semantics operates for alethic modality. It runs into some problems when capturing ethical claims, however. Possible worlds make sense when considering truth-valued statements

capturing states of affairs. A universal state of affairs will give each intelligible proposition a truth-value, and so instantiating modal operators in possible worlds is a good decision method for modal logic. Ethical arguments often do not play by the same rules, however. It may not always be the case, for example, that " $OA \& OB$ " implies " $O(A \& B)$ ," although the equivalent statement is certainly true in modal systems. I will discuss this particular issue in the next chapter, as it is central to one of the more powerful arguments against the consistency of dilemma-producing theories.

Another substantive difference between modal and deontic logics becomes apparent when one examines the "A" in " $\Box A$ " or " $OA$ ." In the case of alethic modality, along with most temporal and belief logics, "A" represents some state of affairs, a formula of propositional or quantificational logic. If this holds for deontic logic, as well, the one should read " $OA$ " as, "It is obligatory to bring it about that A." This will do fine for many moral thinkers, but a number of ethical theories make a point to distinguish actions from their consequences. In these cases, it will be necessary to use action-statements as subschemata for deontic claims. With this in mind, one ought to read " $OA$ " as, "It is obligatory to do A." Georg Henrik von Wright, widely accepted as the founder of contemporary deontic logic, makes use of action-statements rather than states of affairs in his 1951 paper "Deontic Logic" (2). This importantly alters how one constructs formulae of deontic logic. In alethic modality, statements of the form " $\Box\Box A$ " or " $\Diamond\Box A$ " are perfectly admissible. "A holds necessarily" is a proposition just as "A" is, so it can itself hold necessarily or contingently. In von Wright's deontic logic, however, nested operators like this will

not be admitted. "It is obligatory to do A" is not an action-statement. As such, " $OOA$ " would not be well-formed.

It should be noted that in the formal language, we can allow "A" and "B" to stand for specific instances of action, rather than general types of action. So "A" could mean "Michael's saving this particular child from drowning in this particular pond at 10:15 am on June 2<sup>nd</sup>, 2018 in Toledo, Ohio." This sort of paraphrase should uniquely describe the particular action in question, so that formal sentences represent instances of action. Compounds of action-statements can be formed as if they were propositional: " $A \ \& \ \sim B$ ," " $A \supset (B \vee C)$ ," and " $(\sim A \equiv B) \ \& \ (A \vee C)$ " are all well-formed (von Wright, "Deontic Logic" 2-3). This does lead to some tricky situations in interpreting certain statements of the logic, as I mentioned above with " $O(A \ \& \ B)$ ." The formal statement can be read back as normal ("It is obligatory to do A and B jointly"), but it should be noted that doing A and B jointly denotes a particular action, and that action may have a different deontic status than each of its component parts (A and B).

I will follow these conventions for the time being, though I will not give a thorough argument here for my preference for action-statements. Suffice it to say that using action-statements does not fail to capture anything essential for arguments about dilemmas, and so maintaining convention is at least unproblematic. It will occasionally be necessary to include action-statements in formulae of alethic modal logic, which becomes important in some arguments against the existence of dilemmas in standard deontic systems. I will recall this use of action-statements when I discuss those in greater detail, but it has minimal

impact here, and it does not significantly affect the meaning of the alethic statements.

Von Wright's logic laid the groundwork for deontic logic in the late twentieth century and on into the twenty-first. His original system was defined axiomatically. He was writing before the popularity of possible-world semantics, which was injected into deontic logic later. As such, rather than carefully defined systems of possible worlds, the system is given in terms of fundamental laws, or axioms, which serve as foundational formulae from which other statements, called theorems, are derived. These are presented in Rod Girle's text on modality in a more concise and straightforward manner than in the original paper (I am altering the notation slightly for the sake of consistency):

$$(C1): OA \equiv \sim P \sim A$$

$$(C2): PA \vee P \sim A$$

$$(C3): P(A \vee B) \equiv (PA \vee PB)$$

$$(C4): "O(A \vee \sim A)" \text{ and } "\sim P(A \& \sim A)" \text{ are not valid.}$$

$$(C5): \text{If "A" and "B" are logically equivalent, then so are "PA" and "PB"}$$

(Girle 171)

I will give a brief explanation of each axiom before discussing which modal systems are the best fit for deontic logics. (C1) simply states the interdefinability of the operators "O" and "P". (C2) is like a law of excluded middle for morality: Either A or its forbearance must be permissible. This is in keeping with intuition, and I will not seriously challenge it here. It could be argued that (C2) and (C1) together beg the question against a certain type of dilemma in which the same action is both



forbidden and required. I am more interested in this project in contingent dilemmas, where two *different* actions are each forbidden or each required, but an agent is unable to execute both obligations, so I will set aside the worry.

(C3) distributes permission across disjunction, so that “You may do A or B” is equivalent to “You may do A or you may do B.” It should be noted that (C3) is equivalent to “ $O(\sim A \ \& \ \sim B) \equiv O\sim A \ \& \ O\sim B$ ” once one accepts (C1). This particular formulation causes problems in some intuitive cases of moral reasoning. One direction of the biconditional gives the agglomeration principle, which I will argue in the next chapter is too strong a statement to be logically fundamental. For now, I will say that it is at least plausible that (C3) holds, though I do not find it admissible as a logical truth.

(C4) is somewhat strange to have as an axiom, since it does not actually give a formula in the system that can be used to derive theorems. Rather, it states a true fact about the logic, since neither of the stated formulae is derivable from the other axioms. Specifying that the statements must not be valid serves mostly to underscore that they are believed to be contingent, and ought to be discovered by normative argument, rather than included as logically fundamental (von Wright, “Deontic Logic” 10-11). Finally, (C5) simply says that equivalent actions have equivalent deontic status. This causes no problems for the current project, and it would seem quite strange to deny it.

I have given a very brief overview of the basics of deontic logic, focusing on standard models. Before moving on, I will also discuss a few less widely-accepted systems for representing norms. I will not use these much in the arguments to come,

but they warrant consideration, since many feel that standard modal-adapted logic does not adequately capture normative structure. One concern that is actually quite relevant to the topic of this paper is the issue of conflicting obligations. Many moral theories prescribe duties that conflict at least prior to analysis. I will argue of course that certain conflicts of duties are absolute or undischageable, but in many cases of apparent conflict, one obligation takes precedence. Certain logicians attempt to formalize these notions by introducing systems that assign obligations a preference ordering.

Mark A. Brown has developed a system of moral preference ordering in his deontic logic. He sets up his operators to indicate whether a state of affairs is consistent with fulfilling obligations, quantifying over sets of obligations (Brown 121). He uses states of affairs as the contents of his deontic propositions, eschewing von Wright's action-statements. Brown establishes an urgency relation between morally required states of affairs. Sets of possible worlds are assigned "degrees of urgency," defined fairly loosely to allow for various theoretical uses. State of affairs B is at least as urgent as A if and only if every degree of urgency attaching to A attaches to B as well (Brown 122-123). Without getting too deep into the technicalities of the system, one can see that certain states of affairs are described as normatively more important or preferable to others.

The motivations behind a system like Brown's are fairly clear. We have the intuition that some wrongdoings are worse than others. Under most theories, for instance, it is better to lie to a person than to kill them. In either case, one will likely be doing wrong, but it might be a matter of degree. Systems like Brown's are

attractive, though very technically complex. In any case, they fail to do away fully with the problem of conflicting obligations. It may be easy to assign certain obligations greater urgency than others (whether they are formulated as obligations to take an action or to ensure a state of affairs), but many cases are murky. If two obligations are similarly urgent, a method of representing difference in urgency will be of little help. A system to formalize these claims is useful, but it does not eliminate the conceptual difficulties that often arise when comparing moral obligations. While these additions and changes may enrich a logical system's reasoning power, they do little to alleviate the problem of apparent inconsistency for conflicting duties.

I will do well to mention something I passed over before, but which was important in von Wright's development of deontic logic: action logic. Unsatisfied with indicating actions by single variables, von Wright developed a logic of action in a book published a decade after "Deontic Logic" appeared. He saw it as essential that a logic of norms also capture the logical structure of action. Since his early attempts, many other theorists have attempted to schematize language about action, with varying degrees of success. Deontic logician John Horty adapted branching-time semantics from temporal modal logic to capture action. He had in mind the express purpose of developing a richer deontic logic based on agency and decision-making. Without going into further detail, I will remark that this project is extremely important to the wider field of deontic logic, since so much of ethical theory revolves around action. For the purposes of this project, however, I will stick to the simpler syntax I have outlined previously. Systems like Brown's and Horty's will be relevant

to the task of understanding moral dilemmas, but the reader will do fine with only a passing familiarity.

In the next section, I will examine specific arguments that claim to disprove the existence of moral dilemmas using certain standard axioms of deontic logic. These arguments will be illuminating for understanding the logical properties of dilemmas, as well as representing the most cogent reasons for rejecting them. I will refute the arguments by challenging some of the axioms, which are taken for granted in deontic systems. I am relying somewhat on the reader's intuitive understanding of the logical elements at play. The explanation I have given in this chapter will, I hope, have familiarized the reader with the basics of the deontic operators and their semantic interpretations. Alethic modality will play an important role in the coming arguments, as well, though again a passing familiarity will do for my present purposes. These logical systems are complex, but they serve to make analysis of ethical arguments much clearer and more rigorous.

## Chapter 3 - Agglomeration and Ought-Implies-Can

The preceding discussion of deontic logic's formal basis will allow me to present some formal arguments against moral dilemmas. The arguments here are intended to show that if one accepts certain axioms of ethics, one is forced to exclude genuine dilemmas from one's moral theory. I will discuss two main arguments, each of which uses a particular pair of premises seen as basic normative or conceptual principles by many philosophers. I will give reasons for and against each such principle, and in the end reject at least one of the principles needed in each case. These arguments represent the most common and most compelling reasons for taking moral dilemmas to be logically inconsistent. Each argument is based on principles that are at least *prima facie* plausible, and which most thinkers will be inclined to agree with at first glance. These axioms are often taken as obviously true, with little need for defense. The problem with that approach is that moral dilemmas also seem intuitively plausible in many cases, but one cannot accept all these apparently basic principles and also admit dilemmas into one's theory. Since I come down on the side of dilemmas, I will reject some of the principles in question.

The first argument I will examine is the more widely discussed of the two (McConnell 5). It hinges on two potential axioms. Each of these formal axioms represents an ethical (or metaethical) principle that is at least *prima facie* plausible. The first of these is the "agglomeration principle," which states that if one ought to

do A and one ought to do B, then one ought to do A and B (Williams 118). In the formal language, this is:

$$(AG): (OA \ \& \ OB) \supset O(A \ \& \ B)$$

The basic idea is fairly straightforward. Under the agglomeration principle, any pair of individual obligations entails a joint obligation to perform all of the individual actions in question. So if I am obligated to pay my taxes and obligated to refrain from murder, then I am obligated to do both jointly. It is plain to see why this is plausible, especially in cases where the individual obligations are unrelated, or at least do not conflict. Some theorists, including Bernard Williams, have taken issue with agglomeration as a general principle, however. One reason for rejecting it is because of its role in dismissing genuine moral dilemmas (Williams 120). So if one intuitively accepts moral dilemmas, one may be inclined to reject the principle.

The second axiom required for the argument is called the “ought implies can” principle. What it means is that, if one is obligated to do A, one must be able to do A. This principle seems fairly intuitive. At least, one would expect a reasonable norm-issuing authority to abide by it, and there are many cases where intuition would strongly suggest that an agent is not obligated to do the impossible. For example, most people would agree that an agent has an obligation to save a drowning child, given that the agent will not seriously endanger herself by doing so. Consider a case where a child is not drowning, but trapped under a car. The child will certainly die very soon (soon enough that there is no chance of help arriving in time), but the agent is alone and unable to lift the car. It seems fairly obvious that the agent has no duty to take an impossible action, however regrettable it is that she cannot do so.

Formalizing this principle requires a somewhat more careful approach than does the agglomeration principle. Taking “A” and “B” to be sufficiently specific (as I mentioned in the previous chapter), a reasonable formalization for the principle might be:

$$(OIC) OA \supset \Diamond A$$

This relies on a certain interpretation of the modal quantifiers, however. After all, it is not at all logically impossible for Jane to lift the car in my example. This is a case of physical or causal possibility. Of course one will maintain that Jane is not required to lift the car. The advocate of the ought-implies-can principle will say that this case is a legitimate application of the principle. It seems plausible that the principle requires one to interpret “can” to indicate physical possibility. In that case, “ $OA \supset \Diamond A$ ” is a correct paraphrase only when the modal operators are taken to denote physical (i.e. causal) necessity and possibility, as well. It may be possible to construct a logic of ability to capture a more nuanced view of the “can” in ought-implies-can, but I will not give one here. I will take the above as the correct paraphrase, using a generalized version in the argument below. It will do well to take note, however, that all appearances of modal operators will have to correspond to the physical interpretation.

The argument dismissing dilemmas is by *reductio ad absurdum*. Formally, it goes as follows:

- |     |                         |           |
|-----|-------------------------|-----------|
| (1) | $OA$                    | (Premise) |
| (2) | $OB$                    | (Premise) |
| (3) | $\sim \Diamond(A \& B)$ | (Premise) |

- (4)  $(OA \ \& \ OB) \supset O(A \ \& \ B)$  (AG)
- (5)  $(\forall X)(OX \supset \diamond X)$  (OIC)
- (6)  $OA \ \& \ OB$  (From 1 and 2)
- (7)  $O(A \ \& \ B)$  (From 4 and 6)
- (8)  $O(A \ \& \ B) \supset \diamond(A \ \& \ B)$  (An instance of 5)
- (9)  $\diamond(A \ \& \ B)$  (From 7 and 8)

The premises give the definition of a moral dilemma. (1) and (2) state that actions A and B are each individually obligatory. (3) paraphrases the claim that A and B are mutually incompatible actions. Of course, “ $\sim \diamond(A \ \& \ B)$ ” should be taken to indicate that A & B are *contingently* incompatible. This is not only because (OIC) is taken to indicate physical or causal possibility (that is, possibility *given the present state of the world*), but also because of the type of moral dilemma in question. I am not interested in dilemmas where the two horns are logically incompatible. That would be a situation where the very same action was both required and forbidden, or where some two complex actions were mutually contradictory, but both were required. “ $OA \ \& \ O\sim A$ ” and “ $O(A \vee (B \ \& \ C)) \ \& \ O(\sim A \ \& \ \sim B)$ ” are two formal examples of this type of dilemma. These dilemmas are more plainly inconsistent than a dilemma where two actions are contingently incompatible, and indeed they are denied by a pair of fairly indispensable axioms in the logic. Those axioms become relevant for the second argument against contingent dilemmas, so I will leave a fuller discussion of contradictory dilemmas until then. In any case, I am here concerned with dilemmas of the form “ $OA \ \& \ OB$ ” where A and B are factually, rather than logically, incompatible. So the paraphrase “ $\sim \diamond(A \ \& \ B)$ ”, with A and B



sufficiently defined to each capture a specific instance of action must be construed to mean that at time  $t$ , given the actual universal state of affairs, A and B cannot both be performed by the given agent.

The argument as given above is fairly straightforward. By agglomeration, if A and B are each obligatory, then their conjunction is obligatory, as well. The ought-implies-can principle then requires that the agent is able to perform the complex action A & B. However, since the premises describe a dilemma, the agent cannot, in fact, do so. Thus, (3) and (9) stand in contradiction. The opponent of dilemmas will then say that (1) or (2) must be wrong, that is that one or the other of the dilemma's horns is not in fact obligatory. How (and whether) one decides which of the two courses is morally required is a matter that still warrants consideration. Any situation that appears to be a dilemma will involve *prima facie* duties to do both A and B. Thus, in order to deny genuine dilemmas, a theorist will need a way to discern between mere *prima facie* duties and genuine obligations. Mark Brown's comparative obligation-logic provides a formal approach to just this problem, and other logicians have constructed similar systems. Of course, most ethical theories have some way of dismissing certain apparent obligations. Consequentialists use an outcome-calculus and optimize within theory-specific guidelines, and deontologists may say that some duties override or cancel out others.

As effective as the argument from agglomeration and ought-implies-can seems, it has its opponents. I will here offer some responses that should lead the reader to conclude that the argument can at least be plausibly denied. In that case, its principles must not be foundational to a logic of imperatives, since a logical truth

cannot be plausibly denied. Bernard Williams questions the obviousness of the agglomeration principle in a 1965 symposium, which may suggest that one ought not accept it as an axiom (120). That is to say, if agglomeration does hold, it will follow from some more basic principles, and it should not be taken as self-evident or logically trivial. I will follow his argument to some extent, but I will also consider a principle of deontic modality that might guarantee the agglomeration of duties. This principle is the wrong way to approach deontic logic, as I shall explain.

I will begin a *prima facie* counterexample to the agglomeration principle. Consider Tom, who is in love with Catherine. He has promised to marry her, although his parents would prefer he married Elisabeth. To appease his parents, he promises them that he will do so. Assuming one admits promise-keeping as a moral duty, one will say that Tom seems to be facing a dilemma. He ought to keep his promise and marry Catherine, and he ought to keep his promise and marry Elisabeth. However, it would be quite a stretch to say that Tom has a third duty to keep both promises and marry both women. In fact, if one accepts that ought implies can, one will likely have to reject such a duty, assuming Tom lives in a society where polygamy is illegal. Even without appeal to ought-implies-can, one will likely maintain that keeping both promises would in fact be morally wrong, since both women accepted promises of marriage under the implicit assumption of monogamy. It would violate the spirit of both promises if both were kept, so however one approaches the issue, Tom appears to be in a double bind.

Assuming that Tom cannot (or should not) marry both women, the supporter of agglomeration may have a few replies to the example. First, one could say, "Tom

is required to do something he cannot.” Supposing for argument’s sake that ought-implies-can holds, one cannot take this tack. If one has a moral objection to marrying two women, then Tom cannot be obligated to do so if he is obligated not to do the very same. Then perhaps one says, “Tom is not required to marry both women, there exist only two separate duties to keep each individual promise.” Obviously, to say this would be to abandon agglomeration. In the end, if one wants to hold onto both principles, one must say, “One of Tom’s promises (at least) was made in bad faith, so he has no genuine duty to keep it.” This response avoids the apparent difficulty, but it raises some concerns of its own. This line suggests that if one makes a promise one knows is in conflict with another duty, one is not bound by the promise. We might say merely that it is wrong to make such a promise because to do so is deceptive. Once made, the promise holds no weight.

The posited view seems to let the dishonest promise-maker of the proverbial hook for their wrongdoings. Surely individuals are responsible to keep their promises, even and especially when they have promised deceitfully. Suppose that rather than being engaged to Catherine, Tom simply doesn’t like Elisabeth and enjoys the single life. He still promises to his parents that he will marry Elisabeth, though he never intends to keep the promise. Obviously he is bound by it, although he made the promise in full knowledge that he would not keep it. The difference between the cases lies in the fact that when he is engaged to Catherine, Tom has a good reason not to keep his promise. By promising to marry Elisabeth, he promises something that is impossible, assuming he fulfills all of his other obligations. Certainly making this promise is wrong, but to say that it does not generate an

obligation is to misunderstand the situation. Violation of a promise is a wrong act because it is a wrong against an individual. In this case, Tom is beholden to his parents and Elisabeth. He has a duty to these people to fulfill his promise. By violating the promise, he is using them instrumentally and violating their trust. It should be fairly uncontroversial to say one ought not to violate others' trust, and Tom's promise is an instance of this obligation.

The reader may respond: "But of course there are situations where it is permissible to violate another's trust! What if the only way to rescue a bus full of innocent children from a painful death was to break a promise?" This response appeals to quite a correct common sense notion. It may seem that in Tom's case, his promise is similarly overridden. It should first be noted that in the bus-full-of-children case, it hardly seems like a logical truth of imperative logic that one duty should take precedence. However, any plausible *moral* theory will involve a permission to break a promise if doing so saves a bus full of children. In any case, a better response is to point out that the force of the objection comes from the urgency and relative significance of the overriding duty. One obviously feels very strongly that one must prevent the needless and painful deaths of innocent children, especially when one loses little or nothing by saving them.

Tom's case is quite different from the bus-full-of-children example. It is important in Tom's case that his obligations are of equal force. In both cases, he promised to marry someone. The duties are extremely similar. One might point to the difference in the who Tom promised (one promise was to Catherine, the other to Tom's parents), but the example could easily be altered to eliminate any cosmetic

difference between the two promises. The point is that an appeal to the relative force of one duty will not avail the defender of agglomeration. The only remaining objection will be to claim that whichever promise was made in bad faith does not count. This will likely be whichever is made later, although one could appeal to Tom's attitude toward the promise rather than the time in which he makes it.

I stated before that one might balk at letting deceitful promisors off the hook of keeping their promises. This is not to say that every reasonable moral theory will bind a person to every promise they make. The case of the bus full of children shows that it is possible under nearly every theory for one to be released from a promise, even one made legitimately. A more urgent obligation may require one to break a promise, in which case doing so is likely permissible. Likewise, if one is coerced into promising, one is not morally bound to keep the promise. As such, some theories will certainly dismiss at least one of Tom's obligations. The important thing is that at least one plausible theory will *not* discharge one or the other of Tom's promises. Any theory under which Tom is obligated to keep each promise individually but not to marry both women will reject the agglomeration principle. If such a theory is plausible, then agglomeration fails to hold as a conceptual truth for norms in general. Since that is exactly what a deontic axiom is, a plausible non-agglomerative moral theory is evidence that one ought not to admit agglomeration as an axiom.

Of one could reasonably believe a moral theory under which Tom is obligated to marry each of two women, but not both, one must reject agglomeration as an axiom. Discovering this will likely be somewhat surprising, since agglomeration appears to be such an obvious truth. Indeed, the reader may recall that one of von

Wright's axioms of deontic logic implies agglomeration. In fact, agglomeration follows from a particular sort popular model for deontic logic. Deontic modality (the view that deontic logic is really capturing a notion of normative possibility) seems to imply agglomeration. I will argue that a modal system begs the question of agglomeration and explain why one need not admit agglomeration on logical grounds.

In a standard deontic logic, permissions generate possible worlds, and every obligation is fulfillable in every possible world (Girle 170). This upholds the common analogy with modal logic. In a modal system possibility generates worlds and necessity populates them. Two contradictory states of affairs cannot both hold necessarily. If deontic logic mirrors modal logic then obligation mirrors necessity, so all obligatory actions must be jointly fulfillable. Deontic modality is an attractive model for a semantics of deontic logic. Deontic logic was developed in parallel with modal logic, and many systems explicitly use the basic rules and deductive methods for modal logic. In these systems, the analogy to alethic modality hides certain assumptions.

The reader has likely noticed the issue with this system and the agglomeration principle. A deontically possible world is exactly a world in which every obligation is fulfilled in conjunction. That is to say, if one made an exhaustive list of every action required by some source of normative force, those actions would populate a world bound by the normal rules of logic, in particular the law of non-contradiction. If this is the model for a logic of norms then of course  $OA$  and  $OB$  together imply  $O(A \& B)$ . To say " $OA \& OB$ " is just to say, "In every deontically

possible world, A and B are both performed.” “ $O(A \& B)$ ” says just the same thing: Then agglomeration is in fact a biconditional. That biconditional is axiom (C3) of von Wright’s original system for deontic logic. I need not discuss the second direction of the biconditional here. It is fairly trivial to break “ $O(A \& B)$ ” down into its two components, but that approach does nothing for the argument against dilemmas.

The approach of adapting modal semantics to deontic logic begs the question of agglomeration. The world-populating rule does not simply imply agglomeration; the statements are identical. Then my earlier argument against agglomeration should have the same implication for deontic modality. Deontic logic is meant first and foremost to model the logical properties and relations of imperative statements. If cases involving incompatible obligations are logically possible, then any principle banning such cases is not a logical but rather a theoretic principle. Certainly many, if not most theories will demand agglomeration of duties, but at least one plausible theory will not. Any logically impossible set of principles will be implausible, so it follows that plausibility is a sufficient metric by which to reject agglomeration as an axiom.

The flaws in adapting alethic modal semantics for deontic logic will become even clearer when the reader considers what a logic of *prima facie* obligation might look like. One could presumably develop a system to represent and argue about those obligations that arise out of loose or unrestricted application of a moral rule. The only real objection to developing such a logic would be that moral intuitions are often confusing and may not necessarily match onto one another. This is not a real problem for the project of designing such a logic, only for its ability to fully capture ordinary

language. “*Prima facie* obligations” could be taken to mean “obligations obtained from initial application of an ethical principle.” Thus, Kant’s apparent prohibition against lying might collapse when brought up against a stronger-grounded rule. Nonetheless, applying a moral principle is a consistent, reliable process, even if it is done without a fuller analysis of the options available to the agent, so developing a logic of *prima facie* duties should be possible

*Prima facie* dilemmas are present in nearly every moral system. If a *prima facie* obligation-logic is possible then most ethical theories will allow statements in that logic of the form “ $OA \ \& \ OB \ \& \ \sim \Diamond (A \ \& \ B)$ .” It is not unreasonable to suppose that such a logic could maintain the ought-implies-can principle, in which case it would have to do away with agglomeration in order to account for these situations. *Prima facie* obligations are still imperative in form, so they will follow all the rules of a logic of norms. In other words, since the axioms of deontic logic are not specific to ethical imperatives, *prima facie* norms fit the bill. In this way, one sees that there is at least one possible use of deontic logic that does away with agglomeration. Again, one is enough. If the principle does not hold for all possible systems of normative statements and arguments, then it is not an axiom. Since deontic modality would suggest otherwise, one should reject that model as insufficient to represent normative reasoning in at least some cases.

I will now briefly consider reasons for and against the ought-implies-can principle. It has its opponents, though the majority of thinkers are inclined to accept that ought implies can. There are quite a number of good reasons to do so. I will consider what I take to be one of the more powerful objections to the theory, but ultimately dismiss it. If one has followed me thus far with agglomeration, one will



recognize that I need not in fact reject ought-implies-can. In the end, denying the principle does indeed seem to defy the basic logic of imperatives. As such, the weight of the argument falls on the agglomeration of duties, which is only basic to a logical system that take seriously talk of deontically possible worlds.

The intuition toward the ought-implies-can principle comes largely from cases like that of Jane's inability to save a child trapped under a car. In these situations, an action which might be obligatory is not actually so because of physical limitations on the agent's actions. The limitation exists independently of Jane's desire to save the child. Some other cases complicate the intuition, however. Consider the case of Duke, a man who, due mostly to his upbringing, is inclined toward intentional racism (example adapted from Talbot 3). That is to say, he is psychologically incapable of refraining from certain racist actions, though he is aware that his actions are racist, and endorses them (it is assumed that these racist actions are morally wrong). Thus, at time  $t$ , given the actual state of the world, it is not (physically or causally) possible for Duke not to shout a racial slur at a passerby, say. According to ought-implies-can, this means that Duke is not obligated to refrain from his racist actions.

In Duke's case, one would insist that certain moral results follow from his actions, prior to determining whether an obligation exists. Duke likely owes an apology to the direct victims of his racism. He has good reason to provide reparations to those against whom he has discriminated. He and others should try to prevent or mitigate his actions, and others should avoid complicity in his actions. In brief, Duke has *done something wrong*. Talbot posits these factors from intuition,

and claims that the best explanation for this moral residue is genuine wrongdoing. This approach is somewhat suspect for a few reasons, some of which the reader has likely picked up on already. Nonetheless, I find Talbot's tack indicates the best line of argument against the principle. I will discuss Talbot's argument before presenting what I believe are insurmountably powerful objections.

What the argument boils down to is this: Ought-implies-can is best justified by appeal to the action-guiding power of norms. It relies on a narrow reading of action guidance, however. Under a more robust sense of guidance, wrongdoing in cases where the agent could not have done otherwise actually provides a better explanation for certain reciprocal, compensatory, and preventative actions. In Duke's case, dismissing the principle seems plausible. In that situation, most people will maintain that a racist actor has done wrong, even if their wrongdoing is the result of societal factors beyond their control. The action has moral effects identical or at least extremely similar to those resulting from violating an obligation. It seems reasonable, then, to say that although Duke could not have acted otherwise, he is nonetheless guilty of a moral violation. He ought not to perform racist acts, but he does, and he does so intentionally.

Talbot points to Duke's obligations to apologize for his racist behavior and mitigate it where possible, taking these as evidence of residue similar to that of a violated obligation (3). According to Talbot, an obligation to do A at time  $t$  (perhaps with greater detail specifying place, agent, et cetera) can guide action in a number of ways. Obviously, it guides the agent to perform action A at that time. It may guide the agent's actions beforehand as she makes plans for the future. It may guide an

agent to compensate or “approximate the obligatory act” when she fails in an obligation (Talbot 2). It may also guide others to act in order to prevent the agent’s violating her duty, or else to reprimand or demand apology when she does. In short, there are a number of ways an obligation may influence behavior, both in the obligated agent and those around her.

In Duke’s case, if he has an obligation not to make racist remarks despite his compulsion, one can explain how he comes to be obligated to apologize and make amends. Likewise, one might believe he has an obligation to avoid those situations in which he is most inclined to act badly. Both of these secondary actions are part of the obligation’s “robust action guidance” (Talbot 2). Even though it cannot guide Duke not to make racist remarks, since that is beyond his control, his obligation not to do so is intelligible, according to Talbot. He goes on to argue for his position by attempting to categorize how theories will explain the phenomenon of robust action guidance. I will give a brief overview of his discussion on the subject, and then focus on several points where I found Talbot’s example and arguments suspicious.

Talbot argues that theories that reject ought-implies-can lack explanatory power in cases like Duke’s. They must give a separate account of moral residue in cases where ought-implies-can applies. They can do this in one of two ways. Disjunctive theories will claim that certain principles beyond obligation provide robust action guidance when an agent is not in control of her actions. If these same principles apply in cases where an agent is in control, then the theory faces the trouble of overdetermining the moral impacts of actions (Talbot 3.1). When the only difference between two situations is the degree of control an agent has, disjunctive

theories double up on explaining the case of genuine agency, or else open themselves up to disunity by offering two entirely different explanations in very similar cases (Talbot 3.2).

It is not clear to me that Talbot's "doubling up problem" (3.1) is really as much of a problem as he believes. He claims that if two reasons are given for the robust action guidance in cases of agent control then the guidance will be stronger proportionally to the strength of the extra principles (those used to provide robust guidance in non-control situations). This does not follow. It need not be the case that richer explanation for a normative force implies greater power. If the only difference in cases is that in one, the agent has control and in the other she does not then likely there will be no difference in the strength of the robust action guidance. New obligations generated in each case will be the result of whatever principles apply, and the degree of control (and presence of an obligation) will likely alter only one's attitude toward the perpetrator. This is perfectly natural. We do not blame people for what they cannot help.

Talbot offers another solution, replacing disjunctive theories with "elegant" ones (3.3). These theories never appeal to obligations to explain robust action guidance. Rather, they incorporate some other principles, which apply regardless of how much control an agent has over her actions. Talbot identifies certain problems facing elegant theories. Even in situations where it is natural to appeal to obligation (for example, in an apology), one cannot do so on an elegant theory. It is not clear why that is the case. It seems an elegant theory could rely on secondary principles to explain those residues that are relevant when ought-implies-can applies, but allow

obligation to do the work for situations where a certain type of residue only occurs when an agent is in control of her actions. At any rate, it seems that Talbot is not quite explicit enough in just how disjunctive and elegant theories fail to account for robust action guidance. It seems a small bullet to bite in whichever direction one wishes to go.

I believe the strength of Talbot's argument lies largely in his example involving Duke. The problem is that the example is not quite honest. Talbot tries to ride the line between positing a genuinely compulsive or involuntary actor and presenting a character whose actions are condemnable. In choosing a man whose wrong actions are racist, he plays into the reader's natural reaction to distance himself from racism. I imagine that if Duke were seizing violently and physically injuring those around him rather than shouting racial slurs, we would be disinclined to demand apologies or reparations. The defining factor is Duke's endorsement of his own racist actions. This indicates not an obligation to refrain from those actions (when he cannot help it), but a corruption of moral character that the reader finds off-putting at best. The example obscures the intent in this case, and in most other examples of involuntary harms I think we would be less quick to grant Talbot his leaps.

One aspect of the ought-implies-can principle may raise some eyebrows: It would seem that, when one gets oneself into an impossible situation, one might be obligated to do the impossible. This could only apply in cases where one rendered oneself incapable of honoring an obligation. Talbot mentions such an example briefly (2). The notion that one might be required to do something one ought to be

able to do seems powerful, but an alternate explanation is available. If this is a case of robust action guidance, one should be guided not to perform whatever actions gets one in the pickle to begin with. Carrie, who schedules a flight when she has promised to pick up her friend, is responsible for her own failure to keep her promise. The defender of ought-implies-can might simply say that one has an obligation not to cause oneself to do wrong later. Carrie may not be obligated to pick up Mindy *once she is on the plane*, since that obligation would be pointless. Rather, she is obligated not to buy the ticket for that particular time, since it will violate Mindy's trust.

In the end, I believe Talbot's endeavor fails. He attempts to dismiss alternate explanations for robust action guidance, but neither counterargument is particularly strong. His reply to the overguidance objection is somewhat stronger, but he succeeds only in saying that his theory is not the only one with those problems (4). All the standard reasons for accepting ought-implies-can seem to survive the debate. It does indeed mean that obligations at least will provide a certain important sort of narrow action guidance. One may still be puzzled as to why one has no obligation to fly around the world rescuing a child per minute if one can be required to do the impossible. In short, ought-implies-can stands up to scrutiny.

While I have defended the ought-implies-can principle, its role in the argument against dilemmas is moot. We have excellent reasons not to accept agglomeration as a logical truth, and so it is inadmissible as an axiom. In that case, while ought-implies-can seems to be a basic fact about normative logic, dilemma-producing moral theories remain intact, at least up to the point of formal consistency. The next

argument draws on different axioms, again attempting to show that genuine dilemmas imply a theory's inconsistency. That argument will ultimately fail, as well.

## Chapter 4 - Deontic Consistency and Entailment

The argument from agglomeration and ought-implies-can is by far the most popular argument for accusing moral dilemmas of inconsistency (McConnell 4). The appeal likely comes from the axioms' intuitive attractiveness, taken together with its treatment by several widely-read contemporary ethicists (Sinnott-Armstrong 137; Williams 118). The principles of agglomeration and ought-implies can are relatively simple to represent in ordinary language. This makes arguing about them mostly a matter of intuition and reasoning, rather than interpretation of obscure logical forms. I think this last contributes most heavily to the rareness of the second formal argument against dilemmas.

The argument I will present here comes from two principle of deontic logic that live closer to the formal heart of deontic systems. One follows quickly from basic definitions of the operators, and the other attempts to account for logical entailment and incorporate modal necessity. The first principle, sometimes called the "principle of deontic consistency" goes formally as follows:

$$(PDC) \sim(OA \ \& \ O\sim A)$$

In ordinary English, this means that the very same act cannot be both required and forbidden. The defender of dilemmas' consistency will quickly object that this principle begs the question against a certain type of moral dilemma, and so it is inadmissible without proof. The problem for the proponent of "A & ~A" dilemmas is that proof of the axiom is not hard to find. Recall that the operators "P" and "O" are defined thus:



$$(C1) OA \equiv \sim P\sim A$$

Another axiom of the system is this:

$$(C2) PA \vee P\sim A$$

(C1) says simply, "An action A is obligatory iff it is impermissible to refrain from A."

This is a natural interpretation of obligation, and it would defy intuition at the very least to deny it. (C2) says that for every action A, either A or its forbearance must be permissible.

Taken together, (C1) and (C2) clearly imply (PDC). Whichever of "PA" or "P~A" holds will determine that either "~O~A" or "~OA" holds. (C1) seems to be the only interdefinition of the operators that captures ordinary use. Permission is exactly non-forbiddance, or at least non-forbiddance after due consideration by whatever normative authority is permitting and forbidding. In the case of ethics, every action is subject to consideration, so questions of the scope of the norm-authority do not apply. Any refutation of (C1) would have to claim either: (i) ~OA & ~P~A, or (ii) P~A & OA were possible. (i) would mean that there are cases in which A is not obligatory, but it is not permissible to refrain from A. (ii) would mean there are cases in which A is required, but an agent may refrain from A. It should be clear that, assuming morality issues some judgment on every action (i.e. moral principles can be applied in any case), every action must have some deontic status, and of course no action can be required and its forbearance permitted. (C2) follows quickly from this, since a moral norm-authority is universal in scope. There is no board that needs to review its pronouncements. Every action or its negation is permitted, since obligation implies permission.

(PDC) follows from (C1) and (C2), both of which are quite necessary to make much sense of deontic logic at all. The second principle required for the argument against dilemmas is what I will call “the principle of deontic entailment.” Formally, the statement is:

$$(PDE) \quad \Box(A \supset B) \supset (OA \supset OB)$$

What (PDE) says seems simple enough at first glance. If A necessarily implies (entails) B, then A is obligatory only if B is obligatory. In other words, A’s moral requirement entails B’s moral requirement when A entails B. An illustrative example will be helpful here. Suppose I make a promise to visit a friend at a certain time. I am clearly required to keep this promise. Keeping a promise is a straightforward obligation that one has in nearly every situation in which one makes a promise. Keeping this particular promise *entails* my visiting my friend at the appointed time. Of course I am obligated to do so, since otherwise I will break my promise. My promise was just to take this action B, so B’s forbearance clearly violates an imperative. The obligation to keep a promise is logically distinct from the obligation to take any particular promise-fulfilling action, but the one follows directly from the other.

This is an example of logical entailment of actions implying entailment of obligations. The reader will recall from the previous chapter that I there interpreted “ $\diamond$ ” as indicating physical or causal possibility. The distinction between physical and logical necessity now becomes quite important. I will leave a full analysis and counterargument against (PDE) until after I have given the argument against

dilemmas, but I will say for now that the principle comes up against significant difficulties when one is more careful with the modal operator.

The second argument claiming dilemmas' inconsistency goes thus, assuming S5 (whereby all possible worlds are freely accessible) as the modal system:

- |     |   |   |
|-----|---|---|
| (3) | $OA$  | (Premise)   |
| (4) | $OB$  | (Premise)   |
| (3) | $\sim\Diamond(A \& B)$                                | (Premise)   |
| (4) | $\Box\sim(A \& B)$                                    | (From 3 by interdefinition of $\Diamond$ & $\Box$ ) |
| (5) | $\Box(A \supset \sim B)$                              | (From 4 by truth-functional replacement)            |
| (6) | $\Box(A \supset \sim B) \supset (OA \supset O\sim B)$ | (PDE)   |
| (7) | $O\sim B$   | (From 1, 5, and 6)                                  |
| (8) | $\sim(OB \& O\sim B)$                                 | (PDC)   |
| (9) | $OB \& O\sim B$                                       | (From 2 and 7)                                      |

Again, the premises of the argument simply describe a moral dilemma. The truth-functional replacement in (5) will require multiple steps in most rigorous proof systems, but any plausible modal system will preserve necessity across truth-functionally equivalent statements. The use of " $\Diamond$ " to indicate causal possibility in (3) is important. As I mentioned when discussing ability in the previous chapter, this interpretation of the modal operators is useful for capturing notions like an agent's ability, and for discussing what lies in the realm of physical possibility. The problem arises in (6), when the argument makes use of the same modal system to express the principle of deontic entailment. This use assumes (PDE) is meant to apply to instances of causal necessity, wherein one's taking A determines, due to facts about

the world, that one takes action B. The other possible interpretation of (PDE) is that it applies to cases of *logical* entailment. The contradiction between (8) and (9) is clear, and all the inferences are straightforward. It seems again that the best criticisms of this argument will target the axioms.

Above, I offered brief explanations of the axioms, and I defended (PDC) somewhat. The fact that the principle of deontic consistency follows straightforwardly from more fundamental principles at least provides a strong reason to accept it. I will focus my remarks here on (PDE). The principle of deontic entailment is certainly a valuable addition to any system of formalizing ethics. Whenever one obligation gives rise to a related one, chances are good that it does so because of entailment. In my earlier example, I discussed promises. The obligation to keep a promise obviously entails a subsequent obligation to take whatever course of action was promised. An intuitive explanation for that relationship would be that whenever one action entails another, the second is obligatory if the first is. In this case, keeping the promise to visit my friend entails visiting my friend. (PDE) offers a reason for the one duty to give rise to the other.

I will turn shortly to a potential causal account of (PDE). The distinction between logical and causal necessity and possibility may need some further illumination first, though. Both of these concepts fall under the umbrella of alethic modality. As I discussed in my introduction to deontic and modal logic, alethic modality is modality about how states of affairs (or in the present case, actions) may occur possibly or necessarily, or how some states of affairs may necessarily lead to or entail other states of affairs. These (i.e. logical and causal) types of modality

diverge in more than way. On the one hand, an action (or state of affairs) may logically entail another based on how it's phrased. If "A" stands for "the agent in question bandages the man (Robert) who will be robbed by Benjamin" (example from Castañeda 14), then A entails B by first-order logic, where "B" stands for "Benjamin robs Robert." One of course would not say that A caused B to happen in any serious sense. This particular case illuminates a problem with considering (PDE) as addressing logical entailment, which I will discuss later.

Logical modality concerns statements either of actions or states of affairs that have modal status due to their construction. All sentences of the form " $p \vee \sim p$ ," for example, are logically necessary. Causal modality can be taken as identical to physical (or perhaps factual) modality. It will typically incorporate some temporal consideration. The principle is that a state of affairs (or action) A is necessary if, given the actual state of the universe at some point in time  $t$ , A obtains in every possible subsequent universal state of affairs. Logical modality is universal: A logically valid statement holds in every possible world. Logical truths are fact-independent, so no state of affairs will preclude a given logical necessity.

How does this apply to Robert and Benjamin? Obviously it is not logically necessary that Benjamin rob Robert. Rather, the conditional "If I bandage the man whom Benjamin will rob, and that man is Robert, then Benjamin will rob Robert." It is entirely plausible that Benjamin does not rob Robert, in which case the antecedent of this conditional comes out false, since I did not bandage the man Benjamin robbed. I bandaged Robert. In the case of causal worlds, access is temporally ordered. Universal states of affairs hold at different times, and each

instant accesses future states, i.e. possible worlds. So modally-considered statements always apply to future worlds. The question remains of how to interpret the principle of deontic entailment when its formal statement, " $\Box(A \supset B) \supset (OA \supset OB)$ ," is ambiguous.

Consider the following situation: I promise to drive a friend to the bank. Unbeknownst to me, my friend intends to rob the bank. I have no reason to believe this. At least up to my knowledge, I have an obligation to drive my friend to the bank. Suppose that, if I do not drive him, he will not rob the bank. Then it is clear that my driving him causally guarantees that he robs the bank. Assuming I have an obligation to keep my promises (at least when I have no reason to believe doing so will bring about ill effect), I am obligated to drive my friend to the bank. My doing so causes me to take a second action, though not intentionally: I assist my friend in robbing a bank. Clearly, I do not have an obligation to enable my friend's robbery. I likely have an obligation to attempt to prevent it. Here, "A" stands for "I drive my friend to the bank" and "B" for "I enable my friend's robbing the bank." Clearly, the antecedent " $\Box(A \supset B)$ " holds, since my driving guarantees the robbery takes place. Since I have an obligation to take action A, (PDE) would give the result that I have a duty to help my friend rob the bank. Since that is obviously not the case, one has reason to scrutinize the situation, or else to reject (PDE) for cases of causal entailment.

One particular point of this example sticks out. When I claim that I have an obligation to fulfill my promise, this is presumably because I am unaware of the unfortunate consequences of the action (under some theories, my obligation will

withstand the exposure of my friend's larcenous intentions, but it is at least plausible that full knowledge of the relevant facts will eliminate my duty). What deontic facts will change if I know my friend's intentions? Presumably, I will no longer have an obligation to drive him. My promise-keeping obligation may be overridden by a separate obligation to prevent a robbery, at least given that preventing the robbery involves negligible effort on my part. The question for (PDE) is whether my not having an obligation to aid in a robbery is sufficient justification for the claim that I cannot have an obligation to drive my friend to the bank. If this were the case, though, I presumably would not have had the obligation to drive my friend in the original case, according to (PDE).

Perhaps I can only have obligations in cases where my knowledge of the relevant facts is complete. That position seems reasonable, but it has problematic implications. Under such a view, agents will face numerous situations in which they become morally responsible for the unforeseeable results of their actions. If a theory of morality is properly action-guiding, it cannot demand agents act on knowledge they do not possess. This would be something like rejecting ought-implies-can. An obligation to take some action without having a reason is similar to a duty to do the impossible. The obligation is inert. Assuming an actor's ignorance is not willful, one would not consider action taken without full knowledge as blameworthy or warranting rebuke. I think most people have a strong intuition that agents are responsible for foreseeable results but not the unforeseeable results of their actions, but I will offer a clarificatory example.

Suppose Michael is walking along beside a pond when he sees a child struggling in the water. It is clear that if he does not help the child, she will drown. It will be no great burden on him to rescue the child; Michael is a strong swimmer and can easily pull her to shore. Any theory that admits positive obligations (obligations to take, rather than only forbear actions) will require that Michael save the girl. Unbeknownst to Michael, just around the next bend in the road is another pond with five young children drowning. If Michael were aware of these children's plight, he could save them. If he saves the girl in the first pond, he will be too late to rescue the others. The intuition is clear in the first case that Michael must rescue the girl. Were he aware of the other children's situation, he would certainly have an obligation to rescue them. Whenever he has an obligation to rescue the others, it is at least plausible that he has an obligation to rescue the others *instead of* the girl. Plausibility is enough here: If (PDE) is a fundamental axiom of normative considerations, then any reasonable moral theory will be subject to it. There are reasonable moral theories under which saving five lives is preferable to saving one. Suppose Michael has an obligation to rescue the five children in spite of his ignorance. Obviously, he cannot do so without continuing past the drowning girl in the pond. In that case, Michael is not only permitted, but actually obligated to walk past the child drowning in the pond, although he has no reason to believe his actions will lead to his rescuing more children. This is plainly an unacceptable conclusion.

In the above case, it is clear that morality requires Michael to take those actions that he is obligated to take *up to his knowledge*. To return to the bank robbery example, it is now clear that my original obligation to keep my promise was



genuine. The unforeseeable consequences of my action do not have deontic status (assuming my ignorance, like Michael's, is not due to an unreasonable interpretation of those facts available to me). In that case, the supporter of a causally-interpreted (PDE) may simply claim that the modal operator in the principle's formal statement must be taken to apply only up to reasonable knowledge of the situation. This objection seems palatable, but it fails in the end. Suppose Kim is driving down the road. She sees a child in the road and swerves to avoid hitting him. She is aware that in so doing, she will hit a mailbox. She can see the mailbox before she swerves, but of course would prefer destroying property to killing or injuring the child. She clearly has an obligation not to hit the child.

Is it right to say Kim is obligated to hit a person's mailbox? Kim's swerving is purely instrumental to fulfilling her obligation. When she swerves, she is doing so only in order to avoid hitting the child. Her real obligation is not to take some physical action, but to avoid causing a child's death. It is reasonable to say that instrumental actions are morally required when they are necessary antecedents to fulfilling primary obligations. In fact, very few primary obligations deal with first-order acts like turning a steering wheel within a certain time frame. Instrumental obligations are secondary, but genuine moral duties. Kim's hitting the mailbox is not instrumental to her not striking the child, however. According to the logical model for temporal ordering and causal modality, later actions cannot causally determine previous ones. Nor should they be able to. To say so would be to assault common sense. Thus, there is no sense in which Kim's hitting the mailbox, which takes place *after* her obligation not to hit the child is fulfilled, is causally responsible for the

child's safety. Rather, the two share a common cause. Kim's instrumental action of swerving her car just so (an action which is itself obligatory) causally determines both that she fulfills her duty to avoid the child and that she hit the mailbox.

The question that remains is whether secondary duties to perform certain instrumental actions can generate further duties to take causally determined actions (specifically those later in time). It is clear that instrumental duties have some powers of conferral. Kim's obligation to move her hands and feet in a certain way is inherited from her obligation to swerve her car. This inheritance occurs retroactively. That is to say, if an action A is obligatory and an action B is *causally required* to execute A, then B is obligatory. There seems to be no reason to suppose that the same principle would hold of those actions that come about as unintended (but predictable) results of obligatory actions. Presumably, the child in Kim's case will be distraught by the ordeal. If he cries, one would not say that Kim was obligated to see to it that he does so. Rather, this is a result preferable to his death, but by no means morally required.

Kim's example illustrates that obligation can be causally inherited, but only when the inheriting action is causally required to execute the content of the primary obligation. Assuming one accepts causal determination at least at the macroscopic physical level, one must acknowledge that our actions have vastly far-reaching and intricate causal results. To claim that morality prescribes all such results of any obligatory action is to needlessly overpopulate one's ethical ontology. If there is no good reason to believe that actions confer their deontic status forward in time, and

if denying the same does not invalidate any strongly held moral intuitions, one ought to reject the principle.

What becomes of (PDE) on this view? A more discursive semantics for the particular brand of causal modality in question may allow a useful formalization of the principle, but strict temporally ordered causation will not suffice. On the other hand, the present formation of (PDE) may hold when the modal operators are interpreted differently. As I mentioned before, the relationship of promises to their contents is neatly captured by the principle, so it clearly has some value as a foundational concept in ethics. This use does give rise to an important worry about how one ought to formulate actions when using the principle this way. Earlier, I invoked Hector Castañeda's example with Benjamin and Robert. With this case in mind, the reader will likely have a good idea of where the problem arises. If one phrases one's description as, "Arthur bandages the man (Robert) whom Benjamin will Rob," then Benjamin's robbing Robert is entailed, and obligatory when Arthur's ministrations are obligatory. I will echo Castañeda's solution to the quandary, and conclude with a brief description of how moral dilemmas hold up given the rejection of (PDE)'s original interpretation.

In his discussion, Castañeda defines what he calls "acts prescriptively considered" as acts which can be accurately mirrored in the imperative mood or "mandate" form (19-21). In so doing, he details a number of tests whereby linguistic mirroring of statements into the imperative mood illuminates whether the meaning of certain clauses is captured prescriptively or merely describes context. I will not give a full description of each of his tests, but I think an example will clarify his

rather complex approach. Returning to the Benjamin and Robert example, the test goes thus: The statement “Arthur bandages the man (Robert) who will be robbed by Benjamin” entails “Benjamin will rob Robert.” To examine their deontic relationship, however, we must consider each statement as a mandate. The imperative “Arthur, bandage the man (Robert) who will be robbed by Benjamin” does not entail “Benjamin, rob Robert” (Castañeda 21). Our analysis shows that in the case of language considered descriptively, logical entailment holds, and (PDE) would seem to apply if it ever does for logical entailment. Benjamin’s robbing Robert is not an act prescriptively considered, however. The mandate tests reveal this caveat, allowing us to adhere to the intuitive maxim that descriptions of context must not lead us to paradoxes of ethics.

With the primary worry for (PDE)’s application to logical conditionals out of the way, all that remains is to consider what remains of the principle’s value for a system of deontic logic. Consider the case of promises. Obviously, any correct formal paraphrase of “I fulfill my promise to visit my friend on Thursday” will entail “I visit my friend on Thursday.” When we apply Castañeda’s test, we discover that the mandate, “Samuel, fulfill your promise to visit your friend on Thursday” implies the mandate, “Samuel, visit your friend on Thursday.” Thus, the original motivation for the principle of deontic entailment is still satisfied. One can meaningfully claim that holding to commitments implies executing their contents. Other cases where duties generate other duties may arise. In all such situations, however, it is important to distinguish between descriptions of acts with moral status and descriptions of states of affairs providing context for those prescriptively considered acts. The mandate

tests provide a useful mechanism for dismissing those parts of act-descriptions that lack deontic status. Any case in which actions A and B are subject to (PDE)'s legitimate application will clearly involve considering A and B prescriptively, and thus pass Castañeda's test.

In the end, (PDE) survives criticism, but in a somewhat weaker sense than its original presentation. The dilemma actions A and B are each considered prescriptively in their individual statements, but no formulation of either will involve a prescriptive account of the other. This is because the actions are only *contingently* incompatible. A's incompatibility with B is contextual to B's deontic status. "I do A, after which I cannot do B" entails "I do not do B," but the pairing fails the mandate test. This makes sense, since "A & B" dilemmas are more intuitively plausible than "A & ~A" dilemmas exactly because the actions have no intrinsic connection when considered deontically. Likewise, I am unable to execute both A and B through any available course of action, but my forbearance of each is a causal result of my performance of the other, so instrumental inheritance fails, as well. One can have one's deontic cake and eat it, too, since I have preserved most of (PDE)'s usefulness while eliminating its effect on contingent dilemmas.

## Chapter 5 - Conclusion and Implications

I have thus far discussed what are widely regarded as the strongest structural critiques of dilemma-producing moral theories. If these arguments were to succeed, that would imply any theory that tolerated genuine dilemmas had an inconsistent basis. The arguments in the preceding chapters attempt to expose contradictory results of dilemmas. If dilemmas have contradictory results, any set of ethical theorems that accommodate dilemmas will imply the same contradictions. I have argued that we do not have adequate reason to accept the axioms of agglomeration and ought-implies-can as logically fundamental to ethics. That is to say, there may exist logically consistent ethical theories that exclude these axioms.

In the second argument, I acknowledged that both the principle of deontic consistency and the principle of deontic entailment were sound logical analyses of basic moral reasoning. The principle of deontic entailment should not apply to every sort of modal conditional, however. Rather, I argued that the scope of the principle was specifically classical *logical* entailment. After clearing up some possible worries with that interpretation, I concluded that the principle of deontic entailment still gave us much of what we needed in terms of its application to promises and generic versus particular actions. Though causal results do not inherit deontic status, an action may be required or forbidden if it is causally instrumental to fulfilling or violating an obligation. A logical analysis of this inheritance would require a formal logic of ability and instrumentality. It is not within the scope of this project to develop such a system. Regardless, any such system would not categorize horns of

dilemmas as instrumentally related. Rather, forbearing the action required by one horn of the dilemma is a regrettable result of the same actions instrumental to fulfilling the other.

If the reader is at least convinced that it is plausible to deny those axioms I have discussed (or the relevant applications thereof), then she has reason to reject them as axioms, since axioms by definition capture fundamental logical laws. If the logic of imperative ethics does not require these principles, they have their place as potential theorems within specific ethical theories. They should not be taken as foundational to a basic logic of ethics applicable to any theory, which is required to judge the consistency of an entire theory. If one rejects these axioms and accepts that dilemmas are possible, what comes next? The notion of genuine, irresolvable dilemmas goes against certain of our intuitions about moral responsibility and culpability. Questions of blameworthiness arise. One will wish to know just what sorts of moral attitudes one should have to dilemmas and dilemmatic choosers. I will here address these questions, and attempt to provide a satisfactory account of the conceptual status of dilemma-producing theories.

I will begin by giving a more concrete example of just what qualifies as a dilemma. Such a situation must involve two equally powerful obligations, each individually inescapable (at least given the agent's possible courses of action). An agent finds himself in a situation where he must choose one or the other of the two actions, and cannot take both. The actions must not be logically antithetical to one another, but only contingently incompatible. A few of the examples I have discussed

previously may fall under this description, but I will offer one that I believe illuminates the important aspects of dilemmas.

Consider the following: Edward is kidnapped and knocked unconscious. He awakens to find himself in a room with two other people. One is his nephew, the other a total stranger. They are both tied up. Edward finds on his person a gun and a note signed by the kidnapper. The note commands him to kill the stranger, or else the kidnapper will kill Edward's nephew, who is wearing a collar capable of delivering a lethal (but minimally painful) shock. If the collar is removed, it will explode, killing everyone present. Edward has ten minutes to make his choice, after which time the kidnapper will release whomever remains alive. The signature identifies the kidnapper as a notorious criminal. His reputation makes Edward confident that the contents of the note are true and that the criminal will fulfill his promises.

Edward obviously has an obligation not to kill. Under any plausible moral theory, he will also have an obligation to rescue his family members from serious harm or death. I claim that it is at least plausible that neither of these obligations overrides the other in this case. The most likely reasons the reader might give if they think Edward cannot have both obligations are the following: Edward cannot fulfill both obligations, so it does not make sense to suppose that neither can be overridden. This would certainly be true if both the agglomeration principle ( $OA \& OB \supset O(A \& B)$ ) and the ought-implies-can principle were to hold. I have argued that neither is fundamental to the logic of ethics. As long as the reader accepts that a logically consistent moral theory could reject one or the other (or both) of these



principles, the reader cannot maintain that Edward *must* be able to discharge one of his duties in this case. Up to plausibility, theories need not adhere to these principles, and so they need not reject dilemmas. I do not suppose that Edward follows any particular moral theory, in order to consider the ramifications of his situation under theories that admit dilemmas.

The other possible reason to say that Edward cannot have these obligations in conjunction is to say, “If Edward refrains from killing a (presumably innocent) stranger, Edward will bring about his nephew’s death. Since he must refrain from killing the stranger, Edward must bring about his nephew’s death. However, Edward obviously has a duty to rescue his nephew. This is a clear contradiction, since Edward is both required and forbidden to take the same action (namely bringing about his nephew’s death). As the reader will recall from the last chapter, however, I argued against an interpretation principle of deontic entailment whereby one action’s causing another and the first action’s obligatoriness implies the resultant action’s obligatoriness. Given that the principle should only be taken to apply in cases of logical entailment, Edward’s obligation not to kill the stranger does not imply an obligation to bring about his nephew’s death. Were he to refrain from the killing, he would in fact be violating his obligation not to allow his nephew to die.

It can be assumed that either person’s death (the stranger’s or the nephew’s) will have roughly the same effect on the world, in terms of loved ones who grieve and goods that will never be achieved. One might contend that Edward’s obligations to his nephew are stronger than those to strangers, but it is also plausible that Edward’s obligation not to kill is stronger than his obligation to save a life. On

balance, the obligations seem to be at least roughly equally pressing. In short, this case is an example of a genuine moral dilemma. Moreover, this example accurately represents many of the difficulties in dealing with dilemmas. The first point to notice is the tragedy of the situation, and the sense of Edward's being forced into an impossibly difficult choice. In this particular example, a second agent is responsible for creating the dilemma. One could imagine cases where Edward finds himself in a similar predicament, but no person is to blame for his position. Alternatively, there are numerous examples of agents getting themselves into dilemmatic situations. I will discuss just such a case shortly.

Consider Edward's possible courses of action in this case. Assume he has no hope of escaping without playing the kidnapper's game. He has just the two choices, then. He can shoot the stranger, saving his nephew. In that case, he commits a murder, and it does not seem a stretch to say that he wrongs the stranger. I will explore just what it means to wrong an individual shortly. In this case, it coincides with an all-things-considered moral wrong, since Edward has an all-things-considered obligation not to kill, and he violated that obligation. Suppose Edward kills the stranger. Given that this is a true instance of a moral dilemma, Edward has acted wrongly in a certain sense; he has committed a wrong. On the other hand, it seems strange to say that Edward is guilty of anything besides choosing one of two equally abhorrent alternatives given him. If anyone is at fault in this situation, the kidnapper is the obvious choice. Can a dilemma-producing theory fully deflect moral responsibility onto the kidnapper, though? Can it? Let us first consider the other alternative.

Suppose Edward cannot bring himself to kill the stranger. He declares this, and the kidnapper electrocutes Edward's nephew. Again, Edward has wronged another person, this time a member of his family. He has also committed an all-things-considered wrong in allowing his nephew to be killed. In this case, too, it would seem harsh to consider Edward fully culpable for his action. We may even praise the character of a man who is unable to bring himself to kill another. Certainly we would not want to cultivate the opposite instinct. Instead, we regret that the decision had to be made in the first place and seek to place blame for its inception, rather than its execution.

How might Edward feel after taking either of these courses of action? We, as neutral parties may find it obvious that he is not guilty in the same sense as he would be if he killed someone (or let someone die) in other circumstances, but Edward will likely be haunted by his choice, whichever course he ultimately takes. If he kills the stranger, he will then live with the knowledge that he was the direct cause of an innocent person's death. In particular, he may blame himself for being the sort of person who could pull the trigger in that situation. At the very least, this will be a terribly painful memory, perhaps even similar to other memories Edward may have of times he violated (presumably less drastic) moral imperatives. Suppose, on the other hand, that Edward allows his nephew to die. He will then carry with him the knowledge that he was unwilling to do something drastic and wrong to save a loved one. It is not hard to imagine that this memory will also be painful, and likely lead Edward to blame himself for neglecting his duty to his family. This feeling again may even feel comparable in force to the regret one has when one

violates some minor moral rule (donating less to charity than one ought, or lying when there is no real need, for example).

Are Edward's feelings apt? Certainly we do not want to blame him for his choice in the same way we would blame him for a similar action in a non-dilemmatic situation. It seems like a moral theory that admits dilemmas could (and perhaps should) maintain some mechanisms for assigning blame and admonishment for wrongdoing, however. Presumably, such mechanisms will be connected in some way to the violation of moral duties. It clearly cannot be a simple one-to-one relationship, however, since that would entail that violating the duty not to kill in Edward's tragic case would carry the same weight of blame and admonishment as would the same violation in less restrictive situation. I will here discuss what we ought to make of blame in theories that allow dilemmas, along with how it might be different to self-blame rather than assign blame as external parties. I will also touch on certain related issues, including questions of punishment and moral character, especially as they connect to blame and dilemmatic agents.

The first question that arises in considering these subjects is just how to define "blame." As it is understood in most philosophical literature, blame is "a response to moral agents on the basis of their wrong, bad, or otherwise objectionable actions or characters" (Tognazzini and Coates 1). Blame is interpreted variously as an objective judgment of character, especially as it aligns with an agent's own moral views (Haji 197), as an evaluation (against some standard of excellence) of an agent's own qualities "manifested in thought and action" (Watson 231), and as strictly interpersonal judgments, separate from any metaphysical

justification (Hieronymi 115). These are classified sometimes as “cognitive theories of blame,” as they hinge on judgments or evaluations, rather than emotions (Tognazzini and Coates 1.1). Alternative views focus on holding agents responsible by reference to reactive moral emotions (Wallace 18) and adding to a belief that an agent did wrong a desire that they do otherwise (Sher 112). Theories like these are called “emotional” and “conative,” respectively (Tognazzini and Coates 1.2; 1.3).

It would be outside the scope of this project to argue extensively for a particular theory of blame. I will attempt to avoid the tensions between these views where possible. Despite this, I will rely on one particular distinction in the literature on blame. Kenner argues in a 1967 paper against Smart’s strictly cognitive view of blame. In the response, Kenner draws a distinction between use and application of “blame” and “blameworthy” (247-248). “Blame” in its ordinary use expresses an emotional attitude toward a regrettable or frustrating result. It need not imply a moral evaluation (Kenner 239-240). “Blameworthy,” on the other hand is more cognitive. Kenner claims that it is “a simple statement of moral condemnation” (248). I maintain that the term is slightly more complex than this, as “blameworthy” is not merely an expression of the moral undesirability of an action or poor moral character of an agent. It carries with it a certain weight, and may indeed imply a sort of moral grade or a rational desire for the agent to have done otherwise. In short, blameworthiness is closer to a cognitive or conative account of blame, whereas common usage of “blame” is often an expression of causal connection along with a certain negative emotional response.

What bearing do theories of blame have on my example with Edward's dilemma? When we understand what blame is and how it operates, it becomes much easier to determine its application to specific situations. Is blame an apt reaction to wrongdoing in dilemmas? Are dilemmatic agents blameworthy, and if not, would blaming them be in some way unfair? In most cases of wrongdoing, we do not hesitate to assign blame, and we typically consider our blaming attitudes to be justified. That is to say, when I am wronged or violated, I am apt to blame the offending party, and when I believe that wrong had a moral component, I consider my blame well-placed. Blaming someone for her actions seems less appropriate in situations where she was not in full control of those actions, however. Kenner touches somewhat upon this issue when he argues that we can blame a compulsive but incompetent driver for accidents he causes, but we cannot blame a man who has a heart attack while driving (241-242).

Thus, there is a certain sense in which unavoidable actions cannot be subject to blame. If an agent truly has no choice in his action, and his restriction is not the result of some prior negligence or weakness of character, then we are inclined to say that he must not be blamed. Kenner maintains that we expect people to avoid certain dispositions, like compulsively driving cars if one is inept at driving. This view motivates a somewhat stricter requirement to release an agent from blame. One can blame an actor whose involuntary or unavoidable action arose out of a situation that one would normally expect the actor to avoid, but not when the actor was in the situation through no fault of his own. This allows one to explain and

corroborate the intuition that one can blame irresponsible agents when their negligence gets them into situations where they cannot help but cause harm.

This approach will run into problems as we consider more carefully what falls into the realm of reasonable expectations. Are people with diagnosable behavioral disabilities to be blamed for negligence of character? It would seem unfair to say so. In any case, Edward's situation is not quite like those. Edward finds himself forced, through no fault of his own to *choose* a course of action that, under certain plausible moral theories, will be morally wrong. Which choice he makes has little moral significance, since each is wrong, and the two duties are equally pressing. Whichever course he takes, however, he has made some choice. If he kills the stranger, he has elected to do so. A course of action was available to him in which he would not commit that murder. Likewise, if he lets his nephew die, he is electing not to follow a path that would result in his nephew's rescue. The task before us is to make sense of just how Edward is and is not free to choose in this case.

The reader may already have noticed that it would be fairly straightforward to say that Edward was not free to do otherwise than he does if we could formulate his duties and actions disjunctively. If we are to blame him for the wrong act of either killing a stranger or letting his nephew die, then we are clearly in the wrong. Edward could not but take one of those actions, so he could not avoid taking the action consisting of their disjunction. It is not clear that this move would be valid, however. Just as the agglomeration of duties fails to stand against counterexamples, we cannot interpret choices disjunctively. The problem here is plain when one

considers the ramifications of making obligations disjunctive. Suppose I have an obligation to care for my family to the best of my ability and a separate obligation not to murder. If one construes these obligations as disjunctive, it becomes obvious that that approach is untenable. I could fulfill the obligation to take either of those actions by taking one action and not the other. If we claim, as I suggested, that one could interpret the disjunctive obligation as dismissing the individual ones, we must say that that result is correct. But of course this is not the case. I may well have this disjunctive obligation, just as Edward may well be obligated to either save his nephew or the stranger, but each individual obligation remains in force. Thus, the problem of Edward's blameworthiness remains.

Unfortunately for Edward, he finds himself in a peculiar situation where he can choose not to do any particular wrong (and so will be blameworthy when he does violate an obligation), but he cannot help but do *some* wrong (and so may presumably be exonerated of blame). It is unclear what to make of a situation in which someone violates an obligation but does not deserve blame for it. Any sense of "obligation" under which this was true would be quite a weak one. What is at play here then is not excuse from blame entirely, but rather a distinction between two types of blame one might assign. It will be important to be clear in exactly what actions and consequences Edward is responsible for. In this particular case, some blame will land with the kidnapper, as one would expect. I will also address how one ought to interpret this discussion as it applies to cases where a dilemma arises separately from any agents' actions.



In Edward's case, he is responsible for the death of whichever of the victims he chooses not to save. His actions guarantee the person's death, and the victim would not die if Edward did not take that particular action. Whichever action he takes will amount to a killing, either directly killing the stranger or by failing to save his nephew when it was within his power to do so. He is clearly not responsible for bringing it about that *some* obligation is violated, however. In this particular case, the kidnapper is responsible for that. Because of his actions, one of two people will certainly die, and if he acted otherwise, neither would presumably be in any danger.

It is entirely possible that in other examples of dilemmas, no one will be responsible in this way. It is only because Edward's circumstances are the result of someone's actions that any party is blameworthy for this second-order wrong. If there were no kidnapper, we would say that Edward's finding himself in this situation was an unfortunate and tragic matter of circumstance, but his own moral situation would be unchanged. His situation would still constitute a dilemma, so he would still be responsible for his actions as much as in the actual case. He would not fill in the kidnapper's role, however. The question of whose fault it is that he finds himself in such dire straits would have a simple answer: "nobody's." What remains in this investigation is to work out just how blame might be assigned, and how to interpret the different sorts of blame at play for Edward's and the kidnapper's actions.

Hieronymi, following Strawson, locates the assignment of blame in judgments of ill will toward others on the part of the offending party (120). Many moral theories are more concerned with actions and effects themselves than with

will, and so when blame has a moral component, it may well involve judgments and attitudes toward actions rather than states of the will. In Edward's case, it would seem strange to claim that he bore ill will toward either of his fellow kidnapping victims. On Hieronymi's view, Edward would presumably be free of all blame, unless we interpret "ill will" quite loosely. This view is attractive, especially since it exonerates Edward and accounts for the fact that his obligations are owed to the particular people involved. He may have a general obligation not to kill, but the particular wrong he commits under either choice is against an individual who has a certain moral claim against him. However, many moral theories will consider actions that cause harm to be wrong even when unaccompanied by ill will. In such cases, agents may well be blamed who had no particular feelings toward the victims of their actions. In Kenner's case of the incompetent driver, for example, it is safe to assume that the guilty man bears no ill will toward the victims of his crashes. Nonetheless, he is culpable for his negligence and invites blame when he does wrong.

It may seem attractive to say that the difference lies not in an agent's malicious intent, but in the direction of the duties in question. Edward has particular duties to each of his fellow kidnapping victims. The kidnapper has similar obligations, but he may violate an undirected obligation not to bring about more wrongdoing. In that case, one might say that Edward is culpable for wronging whichever person he kills, but not blameworthy in an objective sense. Then the wronged party and any peripheral parties affected will have grounds to blame Edward and seek recourse, but he should perhaps not be judged morally bankrupt

in the same way he would be were he responsible for getting himself into the situation to begin with (I will discuss that possibility in greater detail shortly).

It will be helpful to consider how this sort of blame might work. If the culpable agent were not in a dilemma situation, one would of course be inclined to blame him for his transgressions, regardless of who was wronged. This blame has two components. It involves a judgment that the agent acted in a way that indicates poor moral character, as well as an emotional feeling of indignation on behalf of the wronged party. The first component involves an objective judgment. One blames in this sense when one holds the blamed party responsible for violating some duty in a case where they could have avoided doing so. The second sense of blaming does not suppose an all-things-considered wrong is committed. It consists in a sense of personal connection or empathy with the wronged party, and the feeling of indignation is rooted in the victim's being violated. This emotional reaction may come alongside an indictment of character or call for redress against the culpable party, but it does not constitute a judgment.

Dilemmas are peculiar under this view. If the only difference between these two types of blaming is whether an all-things-considered wrong is committed, then Edward is worthy of both under those theories that hold his case as genuinely dilemmatic. An obligation does not amount to much if its violation does not invite a judgment of morally wrong action. Thus, when Edward violates an obligation by choosing which of his fellow victims to kill or let die, he renders himself blameworthy in both senses I have outlined. He is clearly less worthy of indictment

than one who willfully does wrong, however. What force is at play that mitigates one's blaming Edward?

The missing piece here is sympathy. It is specifically because of the tragic nature of Edward's situation that one is disinclined to blame him fully for his actions. When one considers Edward's situation, one recognizes the difficulty of his choice and the fact that he cannot but do wrong. We do not suppose that a perfect moral agent would do otherwise than Edward does, since neither action is preferable to the other. Thus, one can blame Edward for his choice, whichever that is, without indicting his character or calling for redress on behalf of the wronged party. One still judges that Edward acted wrongly, and whatever his particular action, one maintains that he would avoid certain specific negative effects had he acted otherwise. Nonetheless, one cannot hold an attitude of righteous judgment against Edward because he had no choice but to render himself blameworthy in some respect.

I will turn now to the other party involved: The kidnapper. He carries the weight of the most serious moral violation in this example, but my focus has been on Edward. As it should be: This is a project concerned with moral dilemmas, not kidnappings. The kidnapper is important to the moral narrative in my example, however, because he brings it about that Edward becomes a dilemmatic chooser. Obviously, the kidnapper is wrong for kidnapping his victims and for subjecting them to such severe psychological trauma. What interests me in this case is his responsibility for Edward's actions. Is the kidnapper only to blame for his crimes as a kidnapper and torturer? Or does he share guilt for Edward's choice in the

dilemma? Often, when we imagine a moral dilemma, it is a situation where an agent faces a tragic choice because of some unfortunate circumstance, or else because she got herself into the predicament. I believe that shifting the responsibility to a third party more clearly illuminates how moral culpability operates in these situations, and inform our approach to other types of dilemmas.

We have a strong intuition that the kidnapper is really behind Edward's wrongdoing. There are two possible ways to approach assigning culpability here. First, one could say that Edward's wrongful actions are really the results of the kidnapper's actions. Whichever wrong Edward commits, it may as well be the kidnapper's hand pulling the trigger or the kidnapper's words condemning the nephew to die (in the latter case, the kidnapper is also literally responsible for delivering the lethal shock. I am considering whether he is *also* culpable for Edward's failure to prevent the death). On the other hand, it could be that the kidnapper is culpable for a separate evil. It will be simple in most moral systems to defend an obligation not to force others to violate moral duties. In that case, the kidnapper is not responsible for Edward's actions, but he is responsible for the wrongful act of forcing Edward to take some immoral action.

The first formulation of the kidnapper's culpability quickly runs into problems. The reader will recall the previous chapter where I argued that mere causal results of actions do not necessarily inherit obligatory status from the incipient act. In this case, the kidnapper's actions causally entail Edward's choosing one or the other horn of his dilemma. The immorality of Edward's options need not run backward through time, however. The kidnapper's actions are not instrumental

acts for Edward, so they do not inherit moral status in the same way pulling the gun's trigger is forbidden because shooting the stranger is. It is also important to note that the kidnapper does not *command* Edward to take either course. In that case, assuming Edward took the prescribed action, the kidnapper would be responsible in the same way a general is responsible when his troops carry out his orders. Instead, what the kidnapper does is ensure that Edward will commit *some* wrong act by forcing him into a dilemmatic choice. Whichever act Edward chooses, that particular act is his choice, and not the kidnapper's. However one looks at it, the kidnapper's actions do not guarantee a particular outcome, so he is not strictly causally responsible for whichever harm is done.

The kidnapper's wrongdoing is of the second sort. He is responsible for forcing another person to make some immoral choice. If actions are wrong because they violate the wills of their victims, then this is a particularly perverse thing to do. What the kidnapper has done is not caused one or the other person to die, but rather forced a man to compromise his moral character by himself committing a moral violation. Here the question of culpability is straightforward: The kidnapper's action need not inherit its blameworthiness from Edward's choice; rather, the act itself is a violation. Edward has a right to moral indignation toward the kidnapper for forcing him to do wrong, and the impartial judge has reason to chastise the kidnapper, since he committed a severely wrong act in the face of numerous morally acceptable alternatives. Edward's choice is tragic because he cannot but do wrong. The kidnapper has no such defense. He willfully violated a number of moral

strictures, not the least of which was forcing another person to stain his moral history.

It seems quite straightforward to say that the kidnapper is culpable for his wrongdoing, including the particular action of forcing Edward to make an impossible moral choice. Consider now a related example. Suppose Edward knows the kidnapper's game, and rather than being accosted, he requests to participate. Perhaps he wishes to learn how a psychopath's mind works, or perhaps he simply wants to know whether he has it in him to murder a stranger in order to save a loved one. Suppose also that the kidnapper had already kidnapped the other two people involved, so Edward is not implicated in their capture. Rather, upon learning that his nephew will be involved in one of these games, Edward decides to test himself.

In that case, Edward is largely responsible for forcing himself into an immoral choice. Of course the kidnapper has committed a number of morally abhorrent acts, but the particular crime of forcing Edward to choose either of two unacceptable alternatives falls on Edward's own shoulders. When the kidnapper forced him to choose, the kidnapper invited the blame for causing another person to compromise his own moral character by choosing to do wrong. Now, it is Edward who is forcing himself to compromise his moral fiber.

Is Edward blameworthy for this self-violation? Certainly we wish to blame Edward more severely than in the case where he is kidnapped, but is it correct to say that he has wronged *himself*? To say so would dismiss the strong intuition that he has done a greater wrong by those whom he inevitably harms because of his

actions. We are not concerned with Edward as a victim in this case, but rather with the perversity of choosing to later choose to harm some other person. In this case, we might actually say that the initial choice is instrumental to the dilemmatic violation. This view does not quite capture what is going on, however. We need not suppose Edward *intends* either choice he eventually makes from the outset. He might not even be able to reasonably predict what he will do when the time comes. Rather, he is culpable because he makes a choice ensuring he will later do wrong.

We might now invoke the logical-necessity interpretation of the principle of deontic entailment. If we let “A” stand for “Edward chooses to enter a dilemmatic situation” and “B” for “Edward commits some all-things-considered wrong action,” it is quite clear that A logically entails B. Eventually doing some all-things-considered wrong is central to the definition of entering into a dilemma. Thus the antecedent of (PDE)’s main conditional holds:  $\Box(A \supset B)$ . For this particular case, we will need the contrapositive:  $\Box(\sim B \supset \sim A)$ . Then the relevant instance of (PDE) is this:  $\Box(\sim B \supset \sim A) \supset (O\sim B \supset O\sim A)$ .  $O\sim B$  holds by definition, since B is specified to be all-things-considered morally wrong. Thus,  $O\sim A$  follows. Here we see that A is not forbidden because it compromises Edward’s character. Edward himself is the only person who can perform the action A, and by performing it he would void the very strength of character that appears to be in jeopardy. Rather, A inherits its deontic status because forbearing from A is a logically necessary condition of forbearing from B. Edward is blameworthy in this case for committing a moral violation he had every opportunity to avoid. His choice in the dilemma is no different from the choice he



made when forced into the situation. When he elects to enter the dilemma, however, he violates a distinct but related obligation not to do what entails a further violation.

In analyzing Edward's situation, I have paid special attention to blame. The difficulty of assigning blame to dilemmatic agents is, I believe, one of the greatest conceptual strikes against dilemma-producing theories. We naturally associate blame and wrongdoing, and the former seems to require that the blamed party have some option to act to avoid blameworthiness. In clarifying this matter, I have served a double purpose. Those who are attracted to dilemma-producing theories may have a clearer idea of how to approach the subject in application, and those who rely on blame as an argument against dilemmas' admission into ethics may find themselves in need of a new tack.

Over the course of this project, I have explored what I take to be the most compelling arguments against dilemmas on the grounds of consistency. The intuition that dilemmas cannot be allowed in a consistent system of action-guiding principles has some weight. It is supported by several candidate axioms of deontic logic, which together would indeed show dilemmas to be inconsistent. What I have shown is that these axioms are not in fact fundamental to the project of formalizing ethical claims, and so should only be admitted on independent theoretical grounds, rather than taken as axioms of every admissible deontic logic. When these principles are shown to be fallible, the formal arguments against dilemmas fail. I have now also demonstrated that blame can be understood and used meaningfully within theories that admit dilemmas. Without defending any particular normative position, I will end on this: It may be fruitful to further exploration in ethical theory if theorists

consider the value of accepting dilemmas and weaken the restrictions so often put in place to ensure that obligations be never in conflict.

## References

- Brown, Mark A. "A Logic of Comparative Obligation." *Studia Logica: An International Journal for Symbolic Logic*, vol. 57, no. 1, 1996, pp. 117-137.
- Castaneda, Hector Neri. "Acts, the Logic of Obligation, and Deontic Calculi." *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, vol. 19, no. 1/2, 1968, pp. 13-26. JSTOR, JSTOR.
- Cureton, Adam and Johnson, Robert "Kant's Moral Philosophy." *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), ed. Edward N. Zalta.
- Euripides. *Iphigenia at Aulis*. Translated by George Theodoridis, Bacchicstage, 2007. Accessed at PoetryInTranslation.com.
- Girle, Rod. *Modal Logics and Philosophy*. McGill-Queen's University Press, 2000.
- Haji, I. *Moral Appraisability: Puzzles, Proposals, and Perplexities*. New York: Oxford University Press, 1998.
- Hieronymi, P. "The Force and Fairness of Blame." *Philosophical Perspectives*, vol. 18, no. 1, 2004, pp. 115-148.
- Kant, Immanuel. *Metaphysical Elements of Justice: Part I of the Metaphysics of Morals*. 2<sup>nd</sup> ed., Translated by John Ladd, Hackett, 1999.
- Kenner, L. "On Blaming." *Mind*, vol. 76, no. 302, 1967, pp. 238-249.
- Marcus, Ruth Barcan. "Moral Dilemmas and Consistency." *The Journal of Philosophy*, Vol. 77, No. 3, Mar., 1980, pp. 121-136.
- McConnell, Terrance. "Moral Dilemmas." *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), ed. Edward N. Zalta.
- Plato. *Republic*. Translated by Robin Waterfield, Oxford, 1993.
- Sher, G. *In Praise of Blame*. Oxford: Oxford University Press, 2006.
- Sinnott-Armstrong, Walter. "Moral Dilemmas and 'Ought and Ought Not.'" *Canadian Journal of Philosophy*, vol. 17, no. 1, 1987, pp. 127-139. JSTOR, JSTOR.
- Talbot, Brian. "The Best Argument for 'Ought Implies Can' Is a Better Argument Against 'Ought Implies Can.'" *Ergo*, vol. 3, no. 14, 2016.

- Timmerman, Jens. "Kantian Dilemmas? Moral Conflict in Kant's Ethical Theory." *Archiv für Geschichte der Philosophie*, Vol. 95, Issue 1, 2013, pp. 36-64.
- Timmons, Mark. *Moral Theory: An Introduction*. 2<sup>nd</sup> ed., Rowman & Littlefield, 2013.
- Tognazzini, Neal and Coates, D. Justin "Blame." *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), ed. Edward N. Zalta.
- Von Wright, Georg Henrik. "Deontic Logic." *Mind*, vol. 60, no. 237, 1951, pp. 1-15.
- Wallace, R. J. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press, 1994.
- Watson, G. "Two Faces of Responsibility." *Philosophical Topics*, vol. 24, no. 2, 1996, pp. 227-248.
- Williams, B. A. O., and W. F. Atkinson. "Symposium: Ethical Consistency." *Proceedings of the Aristotelian Society, Supplementary Volumes*, vol. 39, 1965, pp. 103-138. *JSTOR*, *JSTOR*.
- Wolenski, J. "Deontic Logic and Possible Worlds Semantics: A Historical Sketch." *Studia Logica: An International Journal for Symbolic Logic*, vol. 49, no. 2, 1990, pp. 273-282.