Kane on

FREE WILL and DETERMINISM

Introduction

<u>Ch. 1</u>:

The free will problem

- In Kane's terms on pp. 5-6, determinism involves prior "sufficient conditions" for what we do.
 - Possible prior conditions include predestination, the decrees of fate, or the past plus the laws of nature (as in scientific or causal determinism, which is our focus).
 - A sufficient condition necessitates whatever it's sufficient for, so determinism apparently means that we have to make the choices we do, given those prior conditions.
- Ultimately, then, if determinism is true, our choices and hence our actions aren't up to us (i.e. free). They're inevitable.

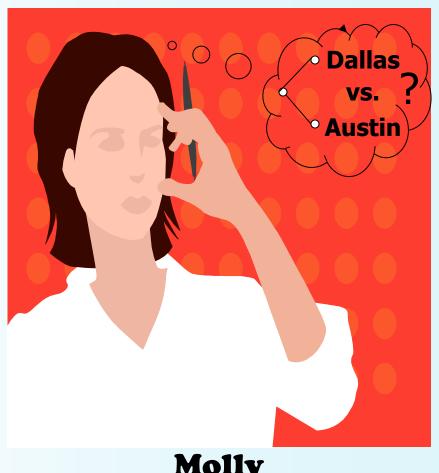
Freedom and responsibility

- "Free will," as Kane interprets it, means more than the "surface" freedom to do or get what we want.
- It also involves control over what we want as opposed to just being manipulated, as (e.g.) in Skinner's Walden Two.
- Freedom is particularly important as a requirement of (moral) responsibility.
 - This involves being blameworthy for our wrong acts and praiseworthy for acts that are particularly good.
 - It wouldn't seem fair to blame someone who ultimately lacked control over what he did (e.g., if his character resulted entirely from childhood abuse, whose effects he couldn't modify later).

Illustrating the conflict

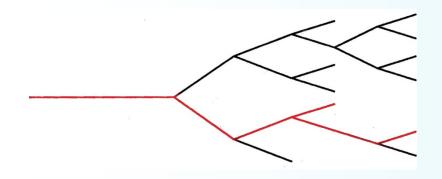
Kane gives the example of Molly, who has to choose which of two law firms to join, a large firm in Dallas or a smaller firm in Austin.

- To think it's worth deliberating about the choice, Molly must believe she has alternative possibilities.
- But determinism implies that actually the choice she makes is inevitable.



Molly

"The Garden of Forking Paths"



- In other words, in order to make choices, an agent like Molly has to assume that she has alternative "forking paths" into the future.
- But if determinism is true, she seems to have only the path she actually takes.
- [Fatalism might allow for forking paths, but they'd all lead to the same place!]

The relevance of modern science

- In the twentieth century the deterministic model of physics has been replaced by quantum indeterminacy.
- However, worries about free will vs. determinism persist for several reasons:
 - Quantum indeterminacy is disputed and might eventually be brought under a larger, deterministic system.
 - Quantum effects are usually insignificant in larger physical systems such as the human brain and body.
 - Quantum events occur by chance, so they wouldn't involve the element of control over action that characterizes free will.
 - Recent developments in sciences other than physics (e.g. neuroscience) tend to favor a deterministic picture of human action.

Considering Compatibilism

Ch. 2:

Arguing for compatibilism

Kane sketches the argument for compatibilism in ch. 2, with the following structure:

- positive: Freedom (or ability; or "can," as in "could have done otherwise) can be understood as conditional on the agent's choices, desires, etc.
- negative: Incompatibilism is untenable, since
 - its "deeper" notion of <u>free will</u> is rationally incoherent
 - its notion of <u>determinism</u> rests on confusion with things that are distinct from it

The hypothetical analysis

- The version of compatibilism favored by modern philosophers. such as Hobbes, Locke, Hume, and Mill, takes the ordinary meaning of freedom as:
 - the power or ability to do what we want, and hence
 - the absence of constraints or impediments to action such as physical restraint, coercion, lack of opportunity, or compulsion.
- In 20th-century terms, this "classical" version of compatibilism can be said to offer a hypothetical or conditional analysis (= "if-then" definition) of "could have done otherwise." Being free to do otherwise [than you actually do] means:
 - you would have done otherwise if you had chosen to
 - you would have chosen otherwise if you had wanted to

Against incompatibilist free will

- Compatibilists go on to question the rational coherency of the "deeper" notion of free will that incompatibilists say they want.
- An incompatibilist alternative to the hypothetical analysis that allows for free will would have to allow for different possible futures ("doing otherwise") with exactly the same past, as in the "garden of forking paths."
- So Molly, in the case presented earlier, would be able to go through the very same deliberation that leads her to prefer the law firm in Dallas and yet choose the Austin firm instead, without any intervening cause to explain why.
- But that would make no sense. It seems that any incompatibilist analysis would yield a rationally incoherent account of deliberation [and hence provide an inadequate basis for responsibility].

Misconceptions of determinism

- Compatibilists also question incompatibilists' understanding of determinism, claiming that it confuses laws of nature or causes with
 - constraint, coercion, compulsion (making us to do something against our will), or
 - control by other agents (as in Walden Two, etc.), or
 - fatalism (our choices have no effect; cf. the "lazy sophism"), or
 - mechanism (humans reduced to the status of robots or amoebas, without conscious reflection or flexible response).
- The point is that determinism works via our will (= desires, choices, etc.). Hume even argued that responsibility requires causation, by our characters and motives (see pp. 18f.), and hence is incompatible with indeterminism.

"Soft" Determinism

- Compatibilism essentially softens the impact of determinism on free will and responsibility, so the combination of compatibilism and determinism is often called "soft" determinism – whereas the combination of incompatibilism with determinism is called "hard," since it has to deny free will and responsibility.
- The issues of determinism and free will turn out to yield four possible positions:
 - soft determinism (compatibilism + determinism, and free will)
 - libertarianism (incompatibilism + free will, so indeterminism)
 - hard determinism (incompatibilism + determinism, so no free will)
 - hard indeterminism (indeterminism + Humean incompatibilism,
 so no free will)

The main positions in matrix form

		Determinism:	
		yes	no
Free will:	<i>yes</i> I:	soft determinism (compatibilist)	libertarianism (incompatibilist)
	no	hard determinism (hard incom	hard indeterminism patibilist)

<u>Ch. 3</u>:

Counterarguments and Rejoinders

- Ch. 3 is structured as back-and-forth between compatibilists and incompatibilists, as follows:
 - Incompatibilists: Van Inwagen's Consequence Argument
 - Compatibilists: The hypothetical analysis renders the argument invalid.
 - Incompatibilists: The hypothetical analysis is faulty, since either
 - it can't handle cases of psychological constraint, or
 - it leads to an infinite regress.

The Consequence Argument

Kane now turns to Van Inwagen's influential argument *against* compatibilism, meant to show that determinism rules out free will. In condensed and clarified form (cf. pp. 23-24):

- 1. We can't now change the past.
- 2. We can't change the laws of nature.
- 3. We can't now change the past and the laws of nature.
- 4. [If determinism is true,] our present acts are the necessary consequences of the past and the laws of nature.
- 5. [If determinism is true,] we can't change the fact that our present acts are the necessary consequences of the past and the laws of nature.
- 6. [If determinism is true,] we can't now change our present acts [as free will requires].

Eluding the argument

Responses to the argument focus on the rule of inference (again condensed a bit) that gets us to step 6 from 3 and 5:

Rule Beta (the "Transfer of Powerlessness Principle"): If no one can change X, or the fact that Y is a necessary consequence of X, then no one can change Y.

- The classical compatibilist can reject Beta by filling in the hypothetical analysis, which yields a different answer for a present act. Thus, in steps 1-6
 - we wouldn't change the past or the laws of nature, even if we now chose/wanted to, but
 - we would change our present acts if we now chose/wanted to.
- On that interpretation, 1-5 would come out true while 6 comes out false, and the argument would be invalid.

Defending the argument

- The incompatibilist can respond in turn by questioning the hypothetical analysis. For instance, it might seem to yield the wrong results for cases of psychologically constraining causes, e.g. those involved in phobias, compulsions, etc.
- Consider McKenna's case of Danielle, who can't tolerate blond Labrador retrievers because of a traumatic childhood experience.
 - Intuitively, it seems that she's unable to touch the blond Lab in a pair of dogs that's presented to her, because she's unable to want to.
 - But the hypothetical analysis would seem to tell us that she is able to touch it, since she would touch it if she wanted to.

Reapplying the analysis

- What if the compatibilist instead tried to capture Danielle's inability to want to touch the blonde lab by applying his analysis to wanting, as well as to acting and choosing?
- He'd then have to say of a normal agent who could but didn't want to touch the dog that she would have wanted to, if she satisfied some further condition (wanted to want to?).
- But questions presumably could be raised again about someone's ability to satisfy that further condition, which would seem to require yet another application of the analysis – and so on ad infinitum.
- This infinite regress would keep the compatibilist from ever fully analyzing ability [so it's a "vicious" regress].

Structure of book

- Introduction (ch. 1)
- Classical compatibilism vs. incompatibilism (chs. 2-3):
 IMPASSE
- Traditional libertarianism (chs. 4-6)IMPASSE
- Hard incompatibilism (ch. 7)PROBLEMS
- New compatibilism (chs. 8-10)OBJECTIONS
- Kane's libertarianism (chs. 11-12)
 ANSWERS TO OBJECTIONS
- Divine predestination (ch. 13)
- Wrap-up (ch. 14)

Defending Libertarianism

Ch. 4:

The Libertarian Dilemma

- Libertarianism (= incompatibilism + free will, so indeterminism) has its own problems.
- It seems that undetermined acts would occur without the sort of control by the agent required for free will and moral responsibility (pp. 33f.; cf. the quote from Hume on pp. 18f.). So either (or both):
 - free will is compatible with determinism, or
 - free will is incompatible with *in*determinism
- This constitutes a dilemma in the logical sense, of a choice between unacceptable alternatives — in this case, unacceptable to a libertarian.

Incompatibilist Mountain

- In Kane's image of "Incompatibilist Mountain" (p. 34) either you can't get up the mountain (to incompatibilism) or you can't get down (to indeterminist free will).
- Getting down is harder than getting up. For Kane it poses "the Descent Problem": explaining how indeterminist free will makes sense.
- The air is thin and cold at the top – by which Kane means that it leads people to invent strange entities (in the case of traditional libertarians) or to contemplate a life without free will (for hard incompatibilists).

[INCOMPATIBILISM]

The Ascent Problem:
Is free will incompatible with determinism?

The Descent Problem:
Can we make sense of and affirm an indeterminist free will?

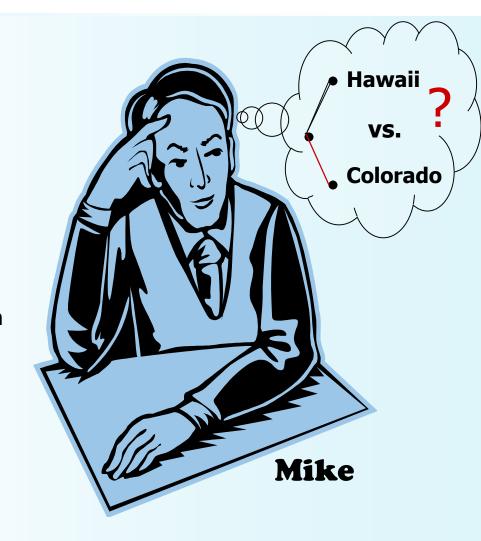
[COMPATIBILISM]

[LIBERTARIANISM]

Kane's illustration

Kane uses a variant of his earlier case of Molly to show how indeterminist free will gives rise to a problem for rational coherency.

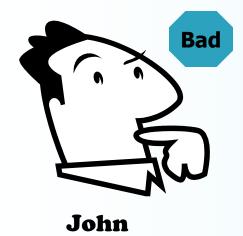
- Here we have Mike, who has to choose between Hawaii and Colorado for his vacation.
- Even if Mike's deliberation merely inclines him toward the choice of Hawaii (rather than necessitating it), his choice of Colorado would still seem to be inexplicable, arbitrary, incoherent.



Random?

- Kane goes on to illustrate several variants of what's sometimes called the problem of "randomness." Various other terms are also used, some of which Kane will later distinguish (e.g., "arbitrariness," as in the "Buridan's ass" case on p. 37), and some of which raise questions particularly about responsibility (e.g. "luck").
- [The idea of random causes of choice isn't just a product of modern quantum physics. In ancient philosophy Epicurus tried unsuccessfully to explain free will in terms of random swerves of atoms falling through the void.]
- In contemporary terms, if Mike's choice were the product of random events such as neurons firing in his brain, that wouldn't give him the kind of *control* required for free action but in fact would seem to be a hindrance to it (see p. 35 on an arm-twitch that interferes with a delicate cut).

The Luck Objection





John*

Some authors pose problems particularly for moral responsibility. Consider, e.g. Mele's case of John and his counterpart in another possible world, John*, who share exactly the same "powers, capacities, states of mind, moral character and the like" up to the moment when each makes a different choice, as allowed by libertarianism:

- John yields to temptation and arrives late to a meeting, whereas John* resists temptation and arrives on time.
- It seems to be just a matter of luck which choice each agent made. But then isn't it unfair to punish John and reward John*?

"The indeterminist condition"

- What stands behind all these cases is the problem of "same past/different futures."
- Kane sums this up in his "indeterminist condition" on libertarian free will (p. 38), ascribing to an agent the ability to act and to act otherwise, given the same past and the laws of nature.
- Libertarians traditionally have tried to accommodate the indeterminist condition by what Kane calls "extra-factor strategies," to be discussed in the next two chapters: introducing something other than a past event, something that isn't subject to natural laws.
- His primary counter-strategy is to show that the resulting views still fail to explain cases like Mike, Molly, and John/John*.

<u>Ch. 5</u>:

Traditional responses

- Traditional libertarian views depend on introducing a further factor, outside the world of science, as an indeterministic cause or type of causation, e.g.
 - the mind, conceived as a separate substance interacting with the body (mind/body dualism, as in Descartes)
 - the noumenal self, as distinct from the phenomenal self (since it's not subject to the laws of nature, or explicable by science and reason, as in Kant)
 - agent-causation, as a different kind of causal relation
 ("immanent," rather than "transeunt" [between events]), as in
 Chisholm, Taylor, and some other twentieth-century figures,
 harking back to Aristotle and Reid).
- However, Kane argues that each of these strategies either is itself subject to a version of the problem of randomness (e.g., p. 42, p. 49) or leaves the extra factor mysterious.

Cartesian dualism

- On Descartes's view, the self is a separate substance, though it interacts causally with the body.
- This is supposed to allow for free will, since past physical circumstances can remain the same while mental activity differs – and potentially changes future physical events, e.g. bodily movements.
- But Kane points out that this isn't sufficient for free will, since if the mind's actions aren't determined by anything (even something about the agent's character, as on Hume's view), they would still seem to be random, etc., along with any acts they cause.

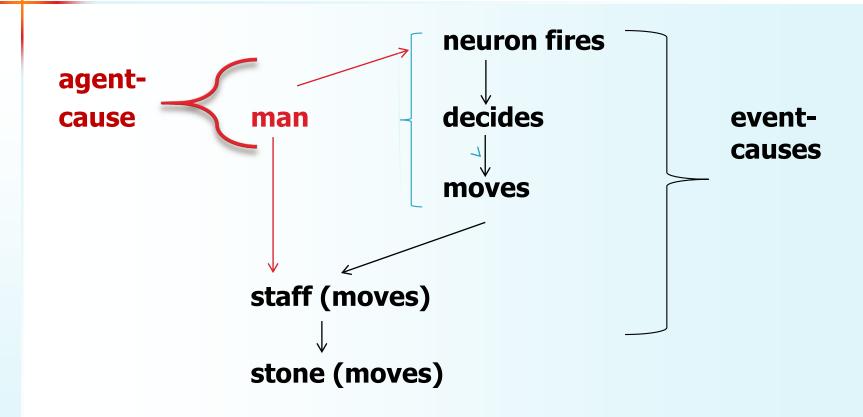
Kantian"double aspect" theory

- On Kant's view, science is limited to studying "phenomena" (or the "phenomenal world") = appearances, organized into objects of experience in accordance with the law of causality, i.e. determinism.
- But Kant also posits "thing-in-themselves" (vs. "as they appear to us"), or "noumena," which we have to suppose stand behind the phenomena. These are unknowable, though we have to believe certain things about them for practical purposes.
- We have to suppose that the noumenal self is free, e.g., in order to think of ourselves as agents. But how this is possible is left mysterious.

Agent-causation

- Though mind/body dualism persists in popular thought, agentcausation is an independent extra-factor approach that's still favored by a number of libertarian philosophers in the current debate (as we'll see in the second half of ch. 6).
- What's "extra" here is not strictly an entity, but rather a different type of causal relationship.
- A free act is said to be caused by the agent, as a substance continuing over time, rather than by prior events, circumstances, states of affairs, etc., including those happening or pertaining to the agent.
- Cf. Chisholm's contrast (p. 46) between a staff moving a stone (event-causation) and a man moving the staff (agent-causation): the latter is a "prime mover unmoved" (cf. Aristotle).

The agent-causal chain



Objections

- Eighteenth-century philosopher Thomas Reid had defended agent-causation as our basic and more familiar causal notion, originating in our earliest experiences of making things happen.
 - But its psychological priority doesn't establish knowledge or the absence of prior event-causes.
 - The agent-causationist view simply stipulates that action isn't caused by prior events. It leaves the proposed mechanism of causation mysterious.
- In any case, can't we just redirect questions of randomness, luck, etc., toward an event involving the new factor: the agent's causing his action?

A regress

- Chisholm responded to the question of randomness by positing a further level of agent-causation, in which the agent causes his agent-causing of a certain act.
- But of course the problem can be raised again at this level: was the agent's causing his agent-causing random? Chisholm was willing to accept an infinite regress here, with repeated applications of agentcausation (but not going back in time).
- Other agent-causationists say instead that agent-causation is nonrandom by its very nature: it amounts to conscious control.
- But this response makes the view doubly stipulative: it just lays down, or stipulates, that agent-causation (1) isn't reducible to event-causation and (2) involves control by the agent.

Ch. 6:

Reasons vs. causes

- A different twentieth-century response to the libertarian dilemma, due to Ginet, claims to do without extra factors in explaining free action and hence is called "simple" indeterminism, though in other ways it's complex.
- Instead of a different type of causation, Ginet allows for a different (but familiar) type of explanation besides causation: explanation by reasons and purposes His view makes two main assumptions:
 - Reasons or purposes are commonly cited in answer to "why?" questions about actions, without necessarily implying anything about causes.
 - Actions are initiated by undetermined volitions (= acts of will), distinguished from mere happenings by the way they feel: their "actish" phenomenal (= experiential) quality. [N.B. Kane at one point treated these as an extra factor, after all.]

Explanation by reasons

- Contemporary causal theorists of action (e.g., Davidson) take desires and beliefs mental states of wanting something and thinking that a certain act is a way to get it or bring it about as reasons that cause (= necessitate) action. [An intentional act, in turn, is distinguished by the fact that it's caused by a desire/belief pair.]
- By contrast to this determinist picture, Ginet holds that reasons and purposes serve to link desires to action, even in the absence of a causal connection.
 - He understands purposes as referring to our desires and as referred to in turn by our intentions [= the mental states we're in when doing (or planning) to do something "on purpose"].
 - An intention has a purpose as its "content" [= what it's about], which can be represented as:

<to bring about some object of desire>

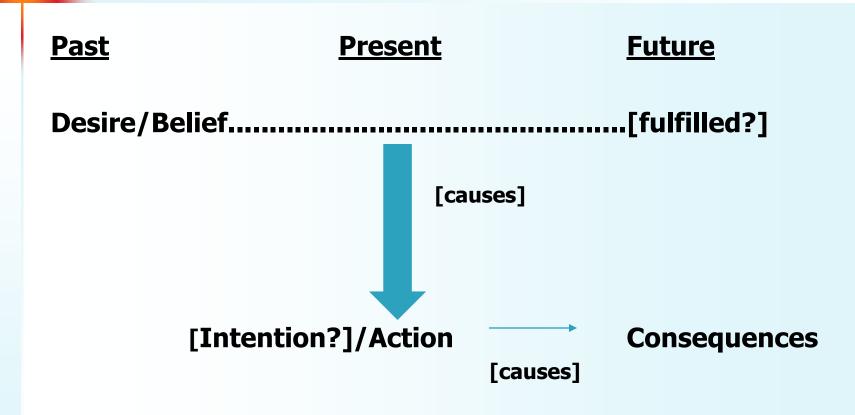
Illustrating the contrast

For example, suppose that the answer to why Mary entered the room is a reason: "to find her keys."

- A causal theorist would explain Mary's action as caused by her desire to get the keys and her belief that they're in the room (or that she can get them by entering the room).
- Ginet would instead explain it by its connection, via her intention, to a purpose that refers to her desire: <to enter the room in order to satisfy the desire to find her keys>.



The causal theorist's picture



Ginet's rough picture

Past Present Future Desire.. [fulfilled?] [refers to] [undetermined] **Intention < Purpose> Volition/Action...** Consequences [achievement [causes] of purpose?]

Problems with the account

- Ginet's noncausal model attempts to answer the problem of randomness, then, by exhibiting a mental state of the agent, an intention, that connects action to desire via purposes.
- However, Ginet tells us nothing about how the intention to act arises or how it gives rise to action. [So the problem of randomness can be raised again at this point; but Kane focuses first on some other objections:]
 - Unconscious purposes might give rise to action without entering into the content of the agent's intentions. An example might be Mary's unconscious desire to wake up her brother by entering the room, ostensibly to find her keys.
 - Also, Ginet says that what initiates action is an uncaused, "actish" mental event called a volition. But Kane notes that we might be deluded about whether something that feels "actish" is really an act, in the sense of being within our control.

Reintroducing agent-causation

- O'Connor thinks we need to refer to an agent-cause in order to explain where an intention comes from [and how it isn't itself a random occurrence], if it isn't causally determined.
- In answer to Goetz's charge of mere stipulation, he claims that agent-causation shows us the structure of choice, distinguishing an action from a mere happening: An agent (A) brings about some event or state of affairs (e).
- O'Connor argues that this structure implies that choices can't be causally determined, since the structure contains no prior event that could be caused (cf. pp. 58-60).
- But Kane points out that the same structure extends to all action, including unfree cases such as coercion or compulsion. So it can't explain what makes an act free.

Mixing agent- and event-causation

- Clarke suggests an indeterministic version of the causal theory of action that needs to be supplemented by an element of agent-causation.
 - The causes of action might just be probabilistic, inclining without necessitating.
 - But in order to assign control to the agent when he makes a choice other than the one his reasons incline him toward (as in Mike's case, or a tie), we need to add in agent-causation as what "tips the balance."
- However [to make a longer story short], Kane and others conclude that Clarke's view has to represent the agent as operating outside the natural causal order of events (p. 63).
- So we'd be back to our original worries about introducing a mysterious extra factor to explain alternative possibilities.

Confining randomness to deliberation

- Dennett and Mele suggested a very different strategy for libertarians to make sense of alternative possibilities: they could accept an element of randomness or luck, but for thought rather than action.
 - Some thoughts that occur to an agent, as he weighs the pros and cons of his options, may be undetermined and hence random.
 - However, his acts or choices aren't themselves random, as long as they're determined by whatever thoughts occur to him (etc.).
- However, Kane says this doesn't really give the agent control over what happens at either stage, of thought (understood as undetermined) or action (assumed to be determined).
- So we've reached another impasse: we still haven't managed to make sense of libertarian free will, as needed to answer Kane's "Descent Problem."

[Mele's later alternative

- Mele has since presented a way of extending his "Modest Libertarianism" (as he later called it) from deliberation to action.
- He thinks of the new view as "Daring Soft Libertarianism" soft, because it needn't be taken as necessary for responsibility, though it's required for the "initiatory power" that some of us (though maybe not all) legitimately value.
- Essentially, the view embraces a degree of randomness as a source of akratic ("weak-willed") action: action in conflict with the agent's judgment of what she ought to do, as in the Molly and Mike cases.
- Over time, however, a rational agent develops increasing control over her action, and her behavior comes to be (more nearly) determined by her "better judgment."]

Doing without Free Will

Ch.7:

Skepticism about free will

- The view that's usually known as "hard determinism" involves denying free will and responsibility.
- Even someone sane, from a basically normal upbringing, who commits a major crime (e.g. Timothy McVeigh), couldn't be considered morally blameworthy on this approach.
- Kane thinks of it as a "cold" view: it leaves you at the top of Incompatibilist Mountain, having established incompatibilism, but unable to make sense of free will.
- A broader contemporary version of the view that's been labeled "hard incompatibilism" (Pereboom) doesn't imply acceptance of determinism.
- Instead, this skeptical position just accepts the Libertarian Dilemma as irresolvable and holds that we'd lack free will and responsibility whether or not determinism is true.

[Galen] Strawson's "Basic Argument"

In simplified form the Basic Argument for hard incompatibilism runs as follows:

- 1. You act as you do because of your character.
- 2. If you were truly responsible for an act, you'd have to be responsible for the character that led to it.
- 3. But if you were truly responsible for your character, you'd have to be responsible for a past act that played a role in making you that way.
- 4. But then step 2 would reapply to that past act.
- 5. But then step 3 would reapply to your character at that time.

...and so on until we get back to early childhood, when you obviously couldn't have been responsible for your character. [The unstated conclusion is that we're not truly responsible for our acts (denying the antecedent of step 2)].

Responding to the Basic Argument

- Though he grants that all premises of the Basic Argument seem plausible, Kane says that a compatibilist or libertarian would object that we can later change our original characters, even if we didn't create them.
- However, Strawson would reply as follows:
 - If the way we change our characters is determined by who we are, then we're not responsible for it.
 - But if the way we change our characters is undetermined by who we are, then it occurs by luck or chance, and again we're not responsible for it.
- In effect, Strawson's argument denies us a role as causa sui
 [= cause of oneself (cf. Nietzsche)] not just originally, but
 also at any later point in our lives, and even partly.

What's lost?

Hard incompatibilists can be thought of as optimists or pessimists, depending on whether they hold that denying free will and moral responsibility would have tolerable or intolerable effects on our view of

- ourselves and our own acts (pride, guilt, etc.; also, motivation, life-hopes, self-image, etc.), and
- other agents, with respect to
 - punishment for wrong acts, and
 - the "reactive attitudes" that underlie personal relationships (love, admiration, gratitude, as well as resentment, blame, etc.; cf. ch. 10 on P.F. Strawson's compatibilism)

Quality of life

- Hard determinist Honderich grants that we'd have to give up the idea that we're ultimately responsible for whether our "life-hopes" are fulfilled, with a loss to our self-image.
- We could no longer be proud of our achievements in quite the same way.
- However, we could retain enough of our life-hopes themselves (success, love, etc.) to make life meaningful.
- We'd still strive to realize them, since we could never know whether they were going to be fulfilled.
- Following Nietzsche, one might even hold that the loss in self-image would make our lives healthier and more honest. We'd be giving up guilt along with (a certain kind of) pride.

Justifying punishment

- In place of retributive punishment, moreover, with its basis in a notion of "desert," we could call upon other motives for punishment, such as:
 - deterrence: prevention of similar acts by the offender or others
 - reform or rehabilitation: modifying the offender's character
- Hard incompatibilist Pereboom also argues that we might reinterpret punishment as analogous to quarantine for serious contagious illness.
- But Kane notes that punishment might then be unfairly meted out to dangerous individuals who haven't actually committed any crime. However, treating punishment as quarantine would rule out subjecting them (or other prisoners) to unnecessarily hard treatment.

Is free will a necessary illusion?

- Pereboom also argues that we'd still place a value on personal love, and we'd retain our admiration for generous or heroic deeds, along with various other [act-centered] "reactive attitudes," e.g. joy in place of gratitude – but not blame or guilt).
- By contrast to these authors, hard determinist Smilansky is a pessimist, maintaining that we need to leave the illusion of free will in place.
 - Without it morality would be undermined, since people wouldn't hold themselves responsible in the same way (by feeling guilty for wrong action, etc.), and
 - The belief in free will is also essential to our own sense of achievement and self-respect.

Reformulating Compatibilism

<u>Ch. 8:</u>

Free will without alternatives

- Nowadays many philosophers understand "free will" just as "whatever sort of freedom is required by responsibility."
- Some, known as "new compatibilists," also challenge the standard assumption that free will involves the ability to do otherwise ("forking paths").
 - They deny what Frankfurt calls the "Principle of Alternative Possibilities" (PAP) for responsibility – and (AP) for free will.
 - Dennett gives some "character-examples" of responsibility without alternatives, e.g. the case of Luther. When Luther said "I can do no other," he clearly didn't mean to avoid taking responsibility for his refusal to recant.
 - But Kane thinks character-examples depend on the assumption that the agent was responsible for earlier choices that created his character. He'd reformulate (PAP) /(AP) to require alternatives only at that earlier time.

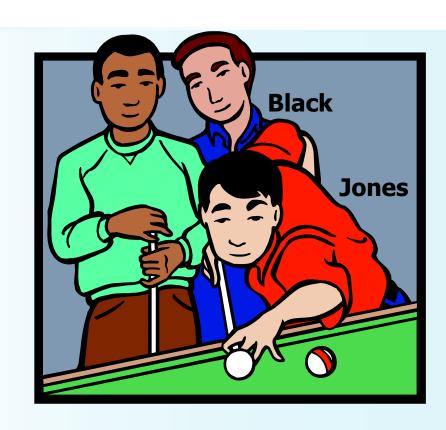
"Frankfurt-type" examples

- A more elaborate kind of example is associated with Harry Frankfurt but was foreshadowed by 17th-century philosopher John Locke and has been modified by later authors to meet objections.
- Locke's case involves a man locked in a room who chooses to stay for some reason of his own (e.g. to continue his conversation) and thus is responsible for the choice, even though in fact he has no alternative to staying in the room.
- However, since he does have an alternative to choosing to stay, some would say that's why he's responsible (for staying, as well as for choosing to stay), if he didn't even try to get out.
- So to rule out even alternative possible choices, Frankfurt sets up a thought-experiment involving mind-control.

A "counterfactual intervener"

In Frankfurt's case, what prevents the agent even from *choosing* to do otherwise is something counterfactual:

- Black is set to make Jones choose A iff Jones doesn't choose A on his own.
- But in fact Jones does choose A on his own, so Black doesn't actually intervene.
- We'd thus hold Jones responsible for choosing and doing A, even though he has no alternative.



Suppose A = throw the game by missing that shot

"Flickers of freedom"

- An influential objection notes that Black's ability to intervene would depend on Jones's providing some prior sign of what he's going to choose to do. So Jones would have had alternative possibilities at that earlier point.
- But John Fischer defends Frankfurt-style cases as follows:
 - If the prior sign is voluntary, we can simply modify the case by shifting Black's (counterfactual) control back to that. [But wouldn't this give rise to a regress?]
 - On the other hand, if the prior sign is involuntary (e.g., blushing), Jones can't be responsible for it. It's only a "flicker" of freedom, in Fischer's term, not robust enough to support responsibility.

The Indeterministic World Objection

- Kane stresses a further objection meant to show that Frankfurt-style cases assume determinism and hence have no force against libertarianism.
 - If Jones's choice were undetermined up to the moment when he chose whether to do A, Black would by then no longer be able to intervene if he failed to choose A on his own, so Jones would have had alternative possibilities.
 - On the other hand, if Black responded to a prior sign and did something earlier to make Jones choose A, then he'd no longer be just a counterfactual intervener, so he, rather than Jones, would be responsible for Jones's doing A.
- Kane therefore thinks the most one can conclude from Frankfurt-style examples is that determinists should be compatibilists (i.e. they should be soft determinists).

"Blockage" cases

- Some authors try to get around the Indeterministic World Objection with cases where Black acts earlier, without a prior sign, but instead of directly making Jones choose A, he just sets up a barrier to alternative choice that won't actually come into play if Jones chooses A on his own.
- A prime example is the Mele/Robb case, where Jones's actual deliberative process leading to his choice of A is indeterministic, but Black has set up another brain process that would preempt it at the moment of choice iff Jones failed to choose A on his own.
- But Kane counters that blocking all alternatives to A would amount to determining Jones's choice in advance, so Jones's process wouldn't really be indeterministic [in context].

New bases needed for free will

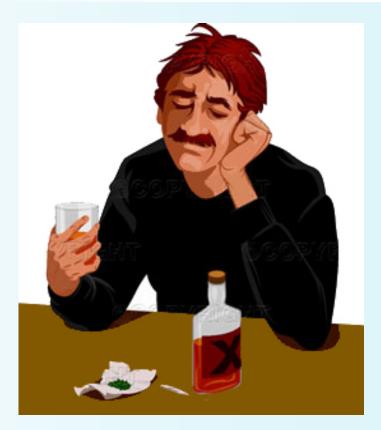
- Even if Frankfurt-type examples don't work against libertarianism, Kane thinks they radically alter the terms of the free will debate, by undermining
 - the Consequence Argument for incompatibilism (covering both libertarianism and hard incompatibilism), and
 - versions of compatibilism based on the hypothetical analysis.
- In general terms, what they tell us is that free will needn't depend on the ability to do otherwise, as all these prior approaches assumed.
- Frankfurt and others go on to give positive accounts of free will, attempting to spell out what it does depend on instead of alternative possibilities.

Ch. 9: A "hierarchical" account

- Frankfurt takes classical compatibilism to capture only freedom of action (Kane's "surface" freedom) by ruling out external impediments or constraints.
- In order to capture freedom of the will, we also need to rule out internal constraints: addictions, compulsions, and other cases where we're not in control of our will.
- Frankfurt attempts to do so by distinguishing between two levels of desire: "second-order" desires are about "firstorder desires" (ordinary desires to act or not); when they're about the desires' effectiveness in action, he calls them second-order volitions [= acts of will].
 - This hierarchical structure underlies the capacity for *reflective* self-evaluation that makes us persons.
 - If we just act impulsively on our first-order desires without reflection, we're "wantons"; complete wantons (with no secondorder desires at all) wouldn't count as persons.

Frankfurt's compatibilism

- On Frankfurt's account, freedom involves the capacity to act on our second-order desires: "having the will [first-order] that one wants [second-order]."
- An "unwilling addict" would be unfree because he has but can't act on a second-order desire not to act on his firstorder desire to take the drug.
- Those of us who can resist acting on our impulses on the basis of reflective selfevaluation [= second-order desires] count as free, whether or not that capacity or its exercise is causally determined.



An unwilling addict

A problem for Frankfurt

- What if we're wantons about our second-order desires?
 - Double gives the case of a cult member who never questions his second-order desire to act on his first-order desire to sacrifice his life if the cult leader asks. Intuitively, he seems unfree.
 - But requiring reflection on higher-order desires would lead to an infinite regress.
- Frankfurt avoids the regress introducing an alternative to third-order reflection, involving:
 - identification: decisive commitment to a second-order desire.
 (But Gary Watson objects to this as arbitrary in its preference for second-order desires, questioning whether they necessarily bear a special relation to the agent.)
 - wholeheartedness: no volitional conflicts (= ambivalence). (But Watson objects that this may be the result of brainwashing or severe conditioning, as in Kane's example of Walden Two.)

A further problem?

- Kane also notes that resolving everyday volitional conflicts (e.g. over what job to take, as in Molly's case, or where to vacation, as in Mike's) is one of the main areas in which we want free will.
- However, on Frankfurt's account we'd be free only after we've resolved such conflicts, since only then are we wholehearted.
 - [But does Frankfurt's requirement of wholeheartedness apply to first-order desires, or only second-order?]
 - In any case, it seems to make a difference to our free will whether we resolve a conflict on our own or as a result of brainwashing, behavioral engineering, etc. But all that matters on Frankfurt's account is how we are at the time of action, not how we got there.

Acting on our values

- Watson's version of the new compatibilist approach to free will harks back to Plato's view of self-control as the rule of Reason over Desire.
- Watson identifies (practical) reason with the agent's values, seen as potentially in conflict with his desires or motives.
 - Free will has to involve the right relation between our valuational and motivational systems, so that we act in accordance with our values.
 - When we act on a desire that goes against the reasons we recognize, we exhibit weakness of will (what the Greeks called akrasia), which Watson takes as unfree.
 - This isn't a question of causal determination, but of compulsive choice, on the model of kleptomania, with Desire exerting influence on action independent of Reason.

Objections to Watson



A kleptomaniac?

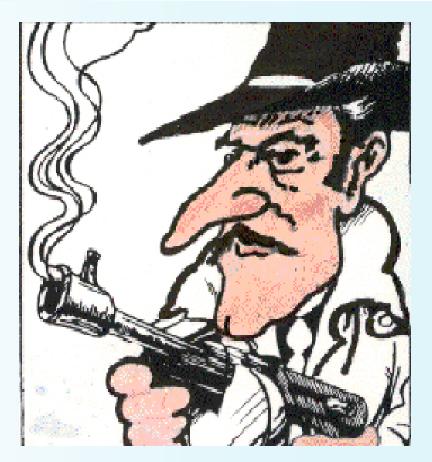
- It's not obvious that weakwilled behavior is always compulsive, in the sense of unfree, however. Don't we hold people responsible for ordinary cases of giving in to temptation?
- Also, as with Frankfurt, Watson's view is subject to problems about behavioral engineering, etc. What about people who are manipulated into always acting on values implanted in them, as in Walden Two?

Requiring "Right Reason"

- Susan Wolf takes Frankfurt's and Watson's views to be "Real Self" or "Deep Self" views insofar as they count us as free when our acts express the self we identify with or value.
- However, she thinks something further is required for free will: the ability to do the right thing for the right reasons, or normative competence.
 - Without this, an agent acting from his Real or Deep Self would count as "insane" in the sense of not knowing the difference between right and wrong.
 - Wolf's own view, which she dubs "the Reason View," results in the asymmetry thesis: that blame but not praise would be undermined by causal determination (understood as making the agent unable to do otherwise). Many people find this implausible.

Objections to Wolf

- Don't we assume in cases of praise that the agent contributed to the formation of his good character – whereas in cases of excuse from blame on grounds of insanity, we assume he couldn't help becoming insane?
- Consider Darth the hit man, who might well have chosen to become what he is and if so would seem to be responsible for what he does. But if, without his knowledge, he were given a drug that made him saintly, he would not seem to be responsible as he would be if he resisted the temptation to revert, once the drug wore off and he had a choice.



Darth the hit man

Ch. 10: Strawsonian compatibilism

Another group of "new compatibilists," following P. F. Strawson, approaches free will and responsibility in terms of our social practice of *holding* people responsible.

- We hold people responsible by feeling certain "reactive attitudes" toward others such as gratitude or resentment (see list, p. 107) for acts seen as expressing good or ill will.
- A "fit" subject of reactive attitudes is an agent who isn't excused or exempt from responsibility according to our practice.
- Our practice is justified as an expression of human needs and concerns, not by metaphysical claims about determinism.
- It would be impossible and irrational to modify our practice in light of determinism, exempting everyone from responsibility, since that would mean giving up our concern with others' good or ill will [as needed for genuinely interpersonal relationships, vs. treating each other as objects rather than persons].

Wallace's version

- Later "reactive attitude theories" attempt to fill out the rationale behind our practice: what it is that justifies excuses and exemptions.
- Wallace appeals to fairness as our reason for not holding certain agents responsible. They're not responsible if either
 - they didn't choose to do wrong, or didn't do wrong deliberately (act out of ill will, in Strawson's terms), so they're excused from responsibility (as with cases of accident, mistake, or nonculpable ignorance; cf. John's failure to pick up Molly, p. 110), or
 - they didn't have the power of reflective self-control, so they're exempt from responsibility (as with very young children, the mentally deficient, the addicted, and the insane).
- Since determinism doesn't imply either condition, Wallace claims it's no threat to responsibility.

Objections to Wallace

- Kane counters that it seems unfair to blame someone unless he's able to do otherwise (contra determinism).
- Wallace would respond in two different ways, corresponding to the distinction between excuses and exemptions:
 - Inability excuses an agent from blame by keeping him from having an obligation (since "ought" implies "can"), so that he isn't really choosing to do wrong. (But Kane gives the case of the elderly man on p. 112 to argue that inability to do otherwise is what explains the difference between his excusable failure to aid the assault victim and his blameworthy failure even to call the police.)
 - Inability exempts an agent from blame only where it's based on the lack of a [general] power -- vs. inability to exercise the power in particular circumstances, as when Judas is "set up" by God to betray Jesus, though he possesses a general capacity for loyalty).

Fischer's "semi-compatibilism"

- Fischer [in effect] proposes a "reactive attitudes" theory of responsibility, but based on Frankfurt-style examples rather than a Strawsonian appeal to social practice.
- Departing somewhat from other new compatibilists [but less than Kane suggests], he grants that
 - freedom [to do otherwise] requires "regulative control," which does involve alternative possibilities, though
 - responsibility [plus a related sense of freedom] requires only "guidance control" – explained in terms of "reasonsresponsiveness": the ability to guide one's behavior by reasons.
- Since determinism would rule out only regulative control, this gives us the basis for a hybrid view called "semicompatibilism," according to which responsibility is compatible with determinism, but freedom [in a sense that Fischer considers irrelevant to practical reasoning] is not.

Illustrating "reasons-responsiveness"

- Fischer (writing with Ravizza) gives the Frankfurt-type case of Mary, whose steering wheel is locked, so that she can only turn left.
- However, she chooses to turn left for independent reasons: in order to get to the mall.
- If the steering wheel were not locked, and she recognized different reasons (e.g. if she thought the mall was on the right), she would have turned right, so she's reasonsresponsive – unlike compulsives, addicts, etc.
- She's therefore responsible –
 e.g., if she hits a pedestrian.



Mary

"Taking" responsibility

- The Strawsonian element of Fischer's view is a further claim that responsibility involves "taking" responsibility: viewing oneself as a fair target of reactive attitudes.
 - Kane explains this as Fischer's rationale for attributing responsibility to some but not all cases that would seem to involve responsiveness to "engineered" reasons, such as predestined Judas and Walden Two.
 - To exclude all such cases, a compatibilist about responsibility apparently would have to appeal to a distinction between control by other agents and determination by impersonal causes. But it's unclear what the rationale would be for that distinction.
- However, couldn't the inhabitants of Walden Two be engineered to take responsibility? Would that make them any more responsible for what they do?

Kane's Libertarianism

Ch. 11:

Ultimate Responsibility (UR)

- Kane reminds us of the two main things we want that incline us toward libertarianism (p. 120; cf. p. 6):
 - 1. <u>alternative possibilities (AP)</u>: to be able to do otherwise than we in fact do
 - 2. <u>ultimate responsibility (UR)</u>: that the sources of our actions be in us, not something else
- AP was stressed by earlier arguments for libertarianism, but Kane now shifts to UR for his answer to the Ascent Problem (getting up Incompatibilist Mountain).
- Later in the chapter, though, Kane extends his argument to show that UR (and hence free will) entails AP.

From UR to incompatibilism

- What UR tells us is that ultimate responsibility for an action requires being responsible for anything that's a sufficient reason, cause, or motive for it.
- So according to UR, if an act results from our character (or motives, etc.), we must have made some past choices that helped to form it but had no sufficient causes themselves.
- To avoid a regress of causes or a denial of free will (as in Galen Strawson's "Basic Argument" in ch. 7), we have to conclude that at least some acts in the agent's history were undetermined.
- So insofar as free will implies UR, it's incompatible with determinism.

"Austin-style" examples

- Some examples suggested by J. L. Austin show that indeterminism isn't sufficient for free will, even if it gives us the ability to do otherwise.
- In their original form these "Austin-style" examples involve doing something accidentally or inadvertently -- e.g., missing a putt, missing the target of an assassination attempt and hitting someone else, pushing the wrong button in a coffee machine -- because of a chance event (e.g. a nervous twitch).
- Even if these acts are undetermined, they're outside the agent's voluntary control and hence don't count as free. (And they do count as acts = things we do.)

Extending UR to the will

- Kane goes on to point out that we also wouldn't be free in an indeterministic world (a "K-world") in which God lets chance play a significant role in action but pre-sets all our reasons, motives, and purposes.
- In Kane's Austin-style example of an assassin trying to kill the prime minister, the assassin's will is set before he acts, but a twitch might make him miss and kill the aide, so his act is undetermined.
- If we alter the case so that he does succeed in killing the prime minister, he's responsible only if he was responsible for setting his will.
- Kane's conclusion is that we need to apply AP and indeterminism specifically to "will-setting" acts on the part of the agent in order to satisfy UR.

The dual regress of free will

- We thus have two separate but related regresses that need to be stopped, according to Kane (see p. 130):
 - a regress of responsibility for action (via voluntary acts as causes of character)
 - a regress of responsibility for motive (via voluntary acts of setting the will = making up one's mind what to do)
- The first regress can stopped by assuming that some of the voluntary acts in the causal chain rendering the agent responsible are undetermined. This is enough to establish incompatibilism (by p. 123).
- But stopping the second regress, as needed to attribute free will to an agent, also requires that there be some point in the causal chain when the agent makes up his mind to act without sufficient prior motives or reasons.

Deriving AP

- Kane argues from free will via UR to AP as follows (p. 129):
 - 1. Free will entails
 - 2. <u>ultimate responsibility</u> (UR) for our wills as well as our actions, which entails indeterminism, particularly in application to
 - 3. <u>will-setting actions</u> at some points in our lives, which in turn entail that some of our actions must satisfy
 - 4. <u>plurality conditions</u>, i.e. that we both act and could act otherwise *voluntarily*, *intentionally* (= knowing what we're doing), and *rationally* (= for reasons), which in turn entails
 - 5. <u>alternative possibilities</u> (AP).

Self-forming acts (SFA's)

- What stops both regresses is a single "will-setting" act, where the will is set in taking action, not before.
- The agent at some point just makes a choice among competing options, though she has no sufficient motive for choosing the particular option she does (in the logical sense of "sufficient"; cf. Buridan's ass, Sartre's example).
- Since undetermined will-setting acts are seen as contributing to the formation of our characters or motives, Kane refers to them as "self-forming acts" (SFA's).
 - UR requires that some of our acts be SFA's.
 - But not all free acts are SFA's; most are consequences of earlier SFA's (as in the character cases discussed earlier, e.g. Luther).

Ch. 12:

Possible scientific bases

- Having ascended "Incompatibilist Mountain," Kane now has to descend – by arguing that responsibility for acts without sufficient causes or motives makes sense.
- His aim is to reconcile libertarian free will with modern science, doing without the mysterious "extra factor" that earlier libertarians appealed to.
- At the outset he combines two elements of contemporary physics to allow for undetermined events in the brain:
 - quantum indeterminism: Determinism fails on the micro-level.
 - chaos (and complexity) theory: Large-scale effects can result from very small changes, e.g. undetermined neuron-firings in the brain can emerge from quantum indeterminacy (with sources in motivational conflict).

Escaping the Libertarian Dilemma

- To answer the charge that indeterminism would take action out of the agent's control, Kane focuses on internal conflict.
- When we're torn between competing options visions of what we should do or become – chaos might be stirred up in the brain, making it sensitive to micro-indeterminacies.
- But either choice would still be rational and voluntary, since we have reasons and motives for it.
 - What we're doing is analogous to blocking out indeterministic "noise" when concentrating on solving a math problem. If we overcome this obstacle and succeed in achieving our aim, we're responsible.
 - This is also illustrated (for non-SFA's) by Austin-style cases, (assuming the agent succeeds – now including Kane's case of an enraged husband intentionally breaking a glass table, where success depends on undetermined factors.

"Doubling"



- Cases of conflict involve "doubling" the indeterministic "noise," so that it's a product of the agent's own will.
- Consider a businesswoman who has to choose between getting to an important meeting and aiding an assault victim.
- Whichever SFA she performs, she'll be succeeding in doing something she was trying to do, for a case of "plural voluntary control,"

Elements of the account

- Kane's account has to provide
 - 1. undetermined acts (with no sufficient cause, motive, reason)
 - 2. plural control (either choice intentional, voluntary, rational)
- He claims to get both from motivational conflict, which
 - 1. generates chaos (and hence indeterminism) in the brain
 - 2. doubles "trying" (and reasons, intentions, etc.)

Questionable elements

- To allow for simultaneously trying to satisfy two conflicting desires, Kane turns to neuroscience for another recent scientific discovery: parallel processing in the brain, whereby we work on two different cognitive tasks at once, e.g. color and shape in vision.
- Agents' endorsement [= acknowledgment] of acts as their own something they were trying to do all along – fills out Kane's reason for assigning responsibility.
- [But does parallel processing apply to conflicting tasks? Why isn't it enough to say, counterfactually, that if the agent had chosen otherwise, she would have been trying to do that alternative act and would later endorse it?
- Perhaps the element of doubled effort is needed to support his account in terms of chaos. But shouldn't Kane go on to ask whether the sense of an act as one's own could be illusory or behaviorally engineered? Cf. his responses to Ginet on "actish" volitions and to Fischer on "taking responsibility" (a term he uses himself as he goes on to answer objections.]

Rebuffing some objections

Kane goes on to consider a number of standard objections to libertarianism, presented in ascending order of significance. The first four are said to rest on errors or overstatements:

- 1. inferring chance from indeterminism. (But indeterminism is consistent with other forms of causation, e.g. probabilistic.)
- 2. separating the act of will from the indeterminism. (But what's undetermined is the agent's effort, choice, etc.)
- 3. taking an element of dependence on luck as ruling out responsibility. (But where success depends partly on luck, as in Kane's modified Austin-style cases, it still may be the agent's doing, if he intends and later endorses it.)
- 4. taking indeterminism as undermining the attribution of choice or control to the agent. (But it doesn't do so completely, if the undetermined choice is embedded in a "self-defining motivational system" realized in her brain.)

Embracing others

- In response to the fourth objection, Kane does allow that indeterminism diminishes the agent's control. (But it comes from the agent's own will, on Kane's account of conflict.)
- He goes on to give two further objections, to which he responds in a similar way, by embracing the charge they lodge against him:
 - 5. that we're not aware of making two competing efforts in SFA's. (But conscious awareness isn't required.)
 - 6. that undetermined choices would be arbitrary. (But agents have good reasons for their undetermined choices, even if not sufficient reasons. Etymologically, "arbitrary" just implies that the decision depends on an exercise of judgment.)

[Beyond SFA's?

- In ch. 12 Kane tries to show two things at once:
 - 1. that undetermined SFA's might result from quantum indeterminism
 - 2. that an element of randomness needn't rule out control by the agent
- If we focus only on 2, leaving 1 to future science, we don't have to restrict free will to acts resulting from characteraltering SFA's. We can extend it to will-setting acts in any situation where we have some reason(s) against the act we choose (if only the effort it takes), so that whichever choice we make would satisfy Kane's plurality conditions, and any randomness would be the result of our own will.]

Ch. 13: Religious disputes

- Kane now turns to an account of the problem of free will in religion, as discussed by Christian philosophers.
- The problem is essentially how to reconcile either predestination or free will with the traits attributed to God (which Kane later notes are a product of Greek philosophy): omniscience, omnipotence, perfect goodness, justice, etc.
 - Jonathan Edwards was a classical compatibilist, defending the Calvinist belief in predestination. But in a religious context that means holding that God created people with evil natures. The problem of free will is intertwined with the problem of evil.
 - Within the Catholic tradition Augustine had defended free will as a gift from God, arguing that an omnipotent God needn't exercise all his powers.

The Foreknowledge Argument

- To defend divine foreknowledge as compatible with free will, Augustine also argued that it doesn't imply causation or necessitation, i.e. predestination.
- However, Kane outlines a version of the Consequence Argument indicating that this doesn't solve the problem.
 - Very briefly: if God knows what we'll do in advance, and his beliefs are necessarily true, and we can't change them, then we can't now change our present acts.
 - Kane goes on to consider four attempts to answer the argument.

Pre-modern proposals (1)

- The <u>eternalist</u> attempt to escape the argument, due to Boethius and defended by Aquinas, holds that God is timeless (outside time), or eternal (exists at all times), experiencing all times as the present.
- So he doesn't really have foreknowledge.
 - But how could He then interact with humans, as in the Bible, assuming He's changeless?
 - In any case, we can't now change what God knows timelessly or at all times either, i.e. the Foreknowledge Argument can be reformulated to apply to timeless or eternal God as well.

Pre-modern proposals (2)

- The <u>Ockhamist</u> proposal, due to William of Ockham (of "Ockham's razor") makes a distinction between "hard" and "soft" facts, with the latter taken as made true by an event at some later time.
- God's foreknowledge of free acts is said to involve knowledge of "soft" facts, which depend for their truth on something we still can change.
 - But this makes out God's "fore"knowledge of our free acts as depending on our later decision to do them.
 - That seems to contradict the view of God as unchanging.

Pre-modern proposals (3)

- A proposal known as <u>Molinist</u>, due to Luis Molina, holds that God has a kind of "middle knowledge" of facts neither necessary nor willed by Him.
- This amounts to knowledge of facts about what each person would do freely if placed in certain conditions = counterfactuals of freedom.
 - [But wouldn't God be able to predict the circumstances we will be in and thus infer what we'll do?]
 - In any case, if nothing necessitates such counterfactuals, all God could know in advance are probabilities, so He wouldn't be omniscient in the full sense.

A modern proposal

- An "Open Theist" position, as in "process" philosophers
 Whitehead and Hartshorne, holds that the future is open,
 and God doesn't know what free agents will do beforehand.
- He still knows everything that has happened and is happening, at all times, which is said to be enough for omniscience, but at any given time He doesn't yet know what hasn't yet happened yet.
 - But then God would no longer be unchanging, or immutable, since he comes to know new things as the world unfolds.
 - Since this qualifies God's perfect nature, this view is rejected by the majority of religious thinkers.

Ch. 14: Wrap-up

- Let's pull things together briefly by listing the five kinds of freedom Kane distinguishes in ch. 14. The first three accommodate a compatibilist view, whereas the last two require libertarianism. Each builds on those preceding.
 - 1. <u>freedom of self-realization</u>: ability to do what you want (classical compatibilists Hobbes, Hume, etc.)
 - 2. <u>freedom of (reflective) self-control</u>: ability to reflect on your reasons and to guide your behavior in light of them (new compatibilists Frankfurt and Watson, Wallace and Fischer)
 - 3. <u>freedom of self-perfection</u>: ability to appreciate the *right* reasons and to guide your behavior in light of them (Wolf)
 - 4. <u>freedom of self-determination</u>: ability to act from a will you were ultimately responsible for forming (Kane's UR).
 - 5. <u>freedom of self-formation</u>: ability to form your will in a way that's undetermined by the past (Kane's SFA's).