

Moral uncertainty – towards a solution?

By [Nick Bostrom](#) · January 1, 2009 4:48 pm · [71 Comments](#) · [« Prev](#) · [Next »](#)

It seems people are overconfident about their moral beliefs. But how should one reason and act if one acknowledges that one is uncertain about morality – not just applied ethics but fundamental moral issues? if you don't know which moral theory is correct?

It doesn't seem you can simply plug your uncertainty into expected utility decision theory and crank the wheel; because many moral theories state that you should not always maximize expected utility.

Even if we limit consideration to consequentialist theories, it still is hard to see how to combine them in the standard decision theoretic framework. For example, suppose you give X% probability to total utilitarianism and (100-X)% to average utilitarianism. Now an action might add 5 utils to total happiness and decrease average happiness by 2 utils. (This could happen, e.g. if you create a new happy person that is less happy than the people who already existed.) Now what do you do, for different values of X?

The problem gets even more complicated if we consider not only consequentialist theories but also deontological theories, contractarian theories, virtue ethics, etc. We might even throw various meta-ethical theories into the stew: error theory, relativism, etc.

I'm working on a paper on this together with my colleague Toby Ord. We have some arguments against a few possible "solutions" that we think don't work. On the positive side we have some tricks that work for a few special cases. But beyond that, the best we have managed so far is a kind of metaphor, which we don't think is literally and exactly correct, and it is a bit under-determined, but it seems to get things roughly right and it might point in the right direction:

The Parliamentary Model. Suppose that you have a set of mutually exclusive moral theories, and that you assign each of these some probability. Now imagine that each of these theories gets to send some number of delegates to The Parliament. The number of delegates each theory gets to send is proportional to the probability of the theory. Then the delegates bargain with one another for support on various issues; and the Parliament reaches a decision by the delegates voting. What you should do is act according to the decisions of this imaginary Parliament. (Actually, we use an extra trick here: we imagine that the delegates act as if the Parliament's decision were a stochastic variable such that the probability of the Parliament taking action A is proportional to the fraction of votes for A. This has the effect of eliminating the artificial 50% threshold that otherwise gives a majority bloc absolute power. Yet – unbeknownst to the delegates – the Parliament always takes whatever action got the most votes: this way we avoid paying the cost of the randomization!)

The idea here is that moral theories get more influence the more probable they are; yet even a relatively weak theory can still get its way on some issues that the theory think are extremely important by sacrificing its influence on other issues that other theories deem more important. For example, suppose you assign 10% probability to total utilitarianism and 90% to moral egoism (just to illustrate the principle). Then the Parliament would mostly take actions that maximize egoistic satisfaction; however it

Subscribe

[All posts](#)
[All comments](#)

Recent Posts

- [Rulesy Folks Push Paternalism](#)
- [Yay Parliaments](#)
- [Why Not Clearer Legitimacy?](#)
- [Our Hidden Motive To Submit](#)
- [Abstract Views Are Coming](#)
- [How To Pick A Quack: Theory](#)
- [The World Forager Elite](#)
- [Elois Ate Your Flying Car](#)
- [Honest Putdowns](#)
- [Specialized Innovation Is Easier](#)

Recent Comments

- [Lord](#)
Standards enable secondary markets by ensuring all products meet minimal conditions. Lacking them, secondary markets could be dangerous. So no counter argument? Even I could do better than that....
[Rulesy Folks Push Paternalism](#) · 13 hours ago
- [Jonathan Starr](#)
Welcome to the party? I want to spew bile at you but instead I will applaud your trait of having enough cash to operate in the first use market. Those of us in the second hand market are grateful...
[Rulesy Folks Push Paternalism](#) · 16 hours ago
- [Jonathan Starr](#)
In that vain I am often buying products in the "so cheep you gotta wonder" range and where I used to shop in stores and visually inspect things or try one and decide if it worked now I just search...

Popular Tags

Academia
Disagreement Ems
Finance
Future
Hypocrisy Law Mating
Medicine Morality
NearFar Personal
Politics
Prediction Markets
Psychology Regulation
Signaling Social Science
Standard Biases Status

Blogroll

- [Andrew Gelman](#)
- [Andy McKenzie](#)
- [Arts & Letters Daily](#)
- [Bayesian Investor Blog](#)
- [Carl Shulman](#)
- [Certain Doubts](#)
- [Choice and Inference](#)
- [Colin Marshall](#)
- [David Friedman](#)
- [Decline Of Scarcity](#)
- [EconLog](#)
- [Eric Barker](#)
- [Eric Drexler](#)
- [Eric Falkenstein](#)
- [Experimental Philosophy](#)
- [Gene Expression](#)
- [George Dvorsky](#)
- [Jess Riedel](#)
- [Katja Grace](#)
- [Kevin Simler](#)
- [Less Wrong](#)
- [Marginal Revolution](#)
- [Matt Ridley](#)
- [Megan McArdle](#)
- [Modeled Behavior](#)
- [Offsetting Behaviour](#)
- [Paul Christiano](#)
- [Philosophy, et cetera](#)
- [Practical Ethics](#)
- [Preposterous Universe](#)
- [Randall Collins](#)
- [Robert Kurzban](#)
- [Robert Wiblin](#)
- [Scott Alexander](#)
- [Seth Roberts](#)
- [Shtetl-Optimized](#)
- [Steven Landsburg](#)
- [Thoughts Arguments and Rants](#)
- [Will Wilkinson](#)

Donate

... be a [charity angel](#).

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- C

But we need to explain why most everyone else goes along

They don't want to be accused of hypocrisy. I want my neighbor to smoke because I don't like the smell and other...

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- [citizen15](#)

I guess the theoretical justification is that there might be some benefit to create some gridlock, I mean to "checking the power of the legislature", without creating a too-strong executive. I...

[Yay Parliaments · 1 day ago](#)

- [RobinHanson](#)

Yes, we should as well ask why there is any space for freedom after all the rules.

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- [Nathan Taylor](#)

I find your arguments convincing, but it begs the question as to why people are so gullible. Or more broadly, why do people innately lean towards blindly following rules (norms)? The answer here...

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- Luke

Rulemaking is a very different process from individual decision-making and vulnerable to very different failures

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- [David Gretzschel](#)

People are irrational, yes, but why does that imply that you can trust people with rulemaking power over other people? That seems extremely dangerous, given how irrational they are. Unless you...

[Rulesy Folks Push](#)
[Paternalism · 1 day ago](#)

- [David Gretzschel](#)

Yes, I think I can agree to that assumption. But do you think that the value of "preventing some, but not all scammy companies

from taking advantage of the naive" to be greater than the compliance...

[Rulesy Folks Push Paternalism](#) · [1 day ago](#)

- [Ronfar](#)
Why do the people on the show "My 600-lb Life" routinely fail to obey the doctor when he tells them if they don't do a specific thing, they're going to be dead within a year? Why do people feel...

[Rulesy Folks Push Paternalism](#) · [1 day ago](#)

- [RobinHanson](#)
But if their situation is as dire as you think, and you are as close as you say, why don't they listen to your advice?

[Rulesy Folks Push Paternalism](#) · [1 day ago](#)

- [Ronfar](#)
"If car companies innovated at the same rate as computer companies, we'd be driving cars that cost \$25 and got 1000 miles per gallon."

"Yes, but would you really want your car to crash twice a day?"

[Elois Ate Your Flying Car](#) · [1 day ago](#)

[Powered by Disqus](#)

would make some concessions to utilitarianism on issues that utilitarianism thinks is especially important. In this example, the person might donate some portion of their income to existential risks research and otherwise live completely selfishly.

I think there might be wisdom in this model. It avoids the dangerous and unstable extremism that would result from letting one's current favorite moral theory completely dictate action, while still allowing the aggressive pursuit of some non-commonsensual high-leverage strategies so long as they don't infringe too much on what other major moral theories deem centrally important.

But maybe somebody here has better ideas or suggestions for improving this model?

Rating: +20 votes

Tagged as: [Epistemology](#), [Morality](#), [Philosophy](#).

Trackback URL:

<https://www.overcomingbias.com/2009/01/moral-uncertainty-towards-a-sol>

71 Comments **Overcoming Bias**   **Pablo Stafforini** ▾

 Recommend 2  Tweet  Share **Sort by Newest** ▾



Join the discussion...

[Load more comments](#)

 [Subscribe](#)  [Add Disqus to your site](#) [Add Disqus](#) [Add](#)

 [Do Not Sell My Data](#)

[Rulesy Folks Push Paternalism](#) · 16 hours ago

- [Jonathan Starr](#)
"one man's red tape is another man's protection." There is another formulation: "Anyone living faster than me is a maniac, and anyone living slower than me is a moron. Me I'm perfect!" As to the...

[Rulesy Folks Push Paternalism](#) · 16 hours ago

- [RobinHanson](#)
Seems you must HATE the fact that people you deal with might have visited a foreign country.

[Rulesy Folks Push Paternalism](#) · 18 hours ago

- [Riothamus](#)
Not Robin, but the answers implied by summary are: 1) Because the executive is selected by the legislature, rather than by the people directly. The parliament picks the prime minister, and the...

[Yay Parliaments](#) · 19 hours ago

- [Lord](#)
No, it only compounds the problem. Now you need to know where everything was bought, who bought it, what its provenance was, was it altered, was it misrepresented. If a gift, bought second hand,...

[Rulesy Folks Push Paternalism](#) · 20 hours ago

- [Tiago](#)
I think you would like the book:)

[Yay Parliaments](#) · 21 hours ago

- [RobinHanson](#)
But it is hard to find externality arguments for most of these policies.

[Rulesy Folks Push Paternalism](#) · 1 day ago

- [Salem](#)
The point about parents and children is interesting, because parents typically bear some of the costs of their children's risk-taking, while getting none of the benefit. So "rational" parents want...