

Rationality and Anticipated Future Desires

Brian Hedden

1 Reflection for Preferences

Our preferences change. If we're rational, how will we take into account anticipated changes in desires or preferences?

A natural answer: When you expect to have some preference in the future, you should adopt that preference now.

Thomas Nagel: 'there is reason to do not only what will promote that for which there is presently a reason, but also that for which it is expected that there *will* be a reason' (*The Possibility of Altruism*, 36)

David Sobel's Reason Transfer Principle: 'If one will later have a reason to get O, then one now has a reason to facilitate the later getting of O.'

Frank Arntzenius: Piaf's maxim: 'a rational person should not be able to foresee that she will regret her decisions.'

John Rawls: 'a rational person does not come to feel an aversion for the foreseen consequences so great that he regrets following the plan he has adopted.'

Preston Greene and Meghan Sullivan's 'Weak No Regrets' principle

Elizabeth Harman on 'I'll be glad I did it reasoning.'

Caspar Hare and Brian Hedden's 'Satisfy Anticipated Desires' principle

Preference Reflection: It is a requirement of rationality that if you believe that you will later prefer *A* to *B*, then you now prefer *A* to *B* (unless you believe that you might be irrational in the future or have lost evidence)

Motivations for Preference Reflection:

1. Characteristic of prudence
2. Parallel to van Fraassen's Reflection Principle for Beliefs
3. A money pump or diachronic Dutch Book argument (compare diachronic Dutch Book argument for Reflection for beliefs, and for Conditionalization)

2 Problems for Preference Reflection

2.1 Bootstrapping and Self-Undermining

Bootstrapping: Suppose that you are deliberating about whether to travel to Argentina or to Brazil. You believe that whatever you do, you will be very glad that you did that thing and not the other. If you go to Argentina, then you will be very glad you went to Argentina rather than Brazil. You will have a preference for having gone to Argentina over having gone to Brazil. And you believe that if you go to Brazil, then you will be glad you chose Brazil over Argentina.

Compare Harman's cochlear implant case

Self-Undermining: Suppose that you believe that whatever you do, you will wish you did the other thing. If you go to Argentina, then you will think fondly of the amazing time you could have had in Brazil and wish you had gone there instead. And if you go to Brazil, you will wish you had gone to Argentina.

2.2 Time-Bias

Bias toward the near

Bias toward the future

The Early Course: You will have 4 hours of painful surgery on Tuesday and 1 hour of painful surgery on Thursday.

The Late Course: You will have no surgery on Tuesday and 3 hours of painful surgery on Thursday.

2.3 Arbitrary Asymmetries

Person/other asymmetry. Parfit (*Reasons and Persons*, 187): 'he may regret that in the past he had his bias towards the near. But this does not show that he must regret having this bias now. A similar claim applies to those who are self-interested. When a self-interested man pays the price imposed on him by the self-interested acts of others, he regrets the fact that these other people are self-interested. He regrets their bias in their own favour. But this does not lead him to regret this bias in himself.'

Past/future asymmetry (compare van Fraassen's principle)

2.4 No Fine-Grained Analogue

Where the A_j form a partition of A , your conditional utility for A given E is defined thus: $U(A | E) = \sum_j U(A_j) \times P(A_j | A \wedge E)$.

Utility Reflection: It is a requirement of rationality that, unless you believe you might be irrational or have lost evidence in the future, then for all A , $U_{now}(A | U_{later} = U) = U(A)$

This entails: $U_{now}(H) = \sum_i U_i(H) \times P(U_{later} = U_i)$

Unfortunately, Utility Reflection faces a version of the problem of interpersonal comparisons of utility.

Utility functions as representations of preferences are unique at most up to positive affine transformations.

Suppose that you are 0.5 confident that you will later have preferences which are representable by utility function U_1 and 0.5 confident that you will later have preferences which are representable by utility function U_2 . Utility Reflection says that your current utility function should be $U_a = 0.5 \times U_1 + 0.5 \times U_2$. But if we change the scale of U_1 , say by multiplying it by 50, to arrive at a different utility function U'_1 which represents the same preferences, Utility Reflection now gives a very different answer about what your current utility function ought to be::

Let $U_a = 0.5 \times U_1 + 0.5 \times U_2$.

Then, $U_a \times 2 = U_1 + U_2$ is a positive affine transformation of U_a .

Let $U_b = 0.5 \times U'_1 + 0.5 \times U_2 = 0.5 \times (50 \times U_1) + 0.5 \times U_2$.

Then, $U_b \times 2 = 50 \times U_1 + U_2$ is a positive affine transformation of U_b .

But $50 \times U_1 + U_2$ is not a positive affine transformation of $U_1 + U_2$.

Possible Solutions

1. Associated introspectable phenomenal quality. But compare Ramsey on degrees of belief
2. The 0-1 rule
3. Sepielli's procedure
4. Extended preferences

3 A Better Principle

Future Welfare Maximization: Other things being equal you ought to maximize your expected future welfare.

But all else might not be equal. Rationality does not require temporal neutrality, nor does it prohibit person neutrality

Advantages

1. Less serious problem of interpersonal/temporal/modal comparisons
2. No bootstrapping or self-undermining
3. Partly mitigates the problem of transformative experience