

Some Varieties of Rationality

We're investigating rational decision making. But there are many different things we might mean when we call a decision “rational”.

There are two basic “inputs” to the process of selecting a rational action for an agent:

- (A) facts about how the agent does, or should, think the world *is*: **beliefs**
- (B) facts about how the agent does, or should, *want to world to be*: **desires/values**

The “or” in “does or should” marks genuine choices in how to think about rationality. For example, here's a very important notion of rationality that focuses on what agents *do* as opposed to what they *should* desire or value:

X's decision to do Y is **instrumentally rational** if *holding fixed* some desires/values/choices of X, Y is among the rational things for X to do.

Consider: What's the most rational way to go about counting the blades of grass in your yard? What's the most rational way for the serial killer to go about killing their victims? When we ask these questions we hold fixed some goal or desire (counting grass, the will to kill indiscriminately), and figure out the best way to achieve that goal or fulfill that desire. Distinguish this instrumental rationality from rationality “full stop”:

If you're going to count the blades of grass in your yard, the *most rational* thing to do is this...
But it's *irrational* to spend your time counting the blades of grass in your yard.

We can draw a similar distinction with beliefs.

X's decision to do Y is **subjectively rational** if *holding fixed* the X's actual evidence Y is among the rational things for X to do.

X's decision to do Y is **objectively rational** if taking into account all the facts Y is among the rational things for X to do.

Consider: Jones has incredibly good evidence that Desert Orchid will win the race, and has a great opportunity to place a bet to this effect. What Jones does not know (and in fact what no one knows) is that Desert Orchid has a fragile femur which lead to injury shortly after the race begins. On which horse *ought* Jones bet?