Mystical Rationality

Isaac Wilhelm

Penultimate draft of Nov 1, 2022.

Please cite the version in De Cruz and De Smedt eds., *Avatar: The Last Airbender and Philosophy*.

The main characters in *Avatar: The Last Airbender* often deliberate in empirically informed ways. When Katara and Sokka start criticizing each other's abilities to perform various tasks around camp, for instance, Aang solves the dispute by carefully observing what everyone's unique talents are. Ultimately, he suggests that they all switch jobs ("The Great Divide"). When Toph, Aang, and Sokka start scamming Fire Nation villagers in order to obtain food, Katara weighs the potential benefits of such activities against the risks involved. Ultimately, she recommends that they stop the scamming ("The Runaway").

The main characters also let mystical, spiritual sources influence their deliberations. When trying to decide whether to kill Fire Lord Ozai, Aang consults with mystical sources—specifically, past Avatars and a lion turtle—to hear their wisdom ("Sozin's Comet, Part 2: The Old Masters"). Though a firebender by birth, Iroh studies all four

nation's spiritual bending practices in order to learn the defensive art of redirecting lightning ("Bitter Work").

In this chapter, we will explore some ways in which reasoning based on mysticism can be rational, focusing on the episode "The Fortuneteller," in which Aang, Katara, and Sokka save a village from a volcano eruption. In this episode, Sokka advocates a purely empirical approach to reasoning. The villagers, however, believe that no source of knowledge is more reliable than Aunt Wu, the local fortuneteller. At several points in the episode, Sokka claims that the villagers' reliance on Aunt Wu is irrational. The villagers claim otherwise: since Aunt Wu has never led them astray, it is rational to rely on her.

As we will see, the villagers are right. Their approach to reasoning—based on Aunt Wu's fortunetelling—is more rational than it initially seems to be. Roughly put, despite their belief in the mystical, the villagers behave in accord with a standard philosophical theory of rational decision-making, a theory of what makes some decisions rational and others not. So Sokka is wrong to claim that the villagers' reliance on Aunt Wu is irrational. The villagers behave quite rationally.

Sokka, the Villagers, and Aunt Wu

In "The Fortuneteller," Sokka finds the villagers' behaviors frustrating. He claims that believing Aunt Wu's predictions is irrational: there is no good reason for thinking that Aunt Wu's predictions will come true. The villagers disagree: they strongly believe in the accuracy of Aunt Wu's predictions. What Sokka finds most frustrating, however, is that Aunt Wu's predictions *do* come true – but only because the villagers believe whatever Aunt Wu predicts.

In particular, the following pattern appears throughout “The Fortuneteller.” Because the villagers believe Aunt Wu’s predictions, they behave in certain ways. Those behaviors bring about the very things that Aunt Wu predicted would happen. Consequently, the villagers take Aunt Wu’s predictions to be confirmed. Of course, in a sense, the villagers are right: Aunt Wu’s predictions are constantly coming true. But the villagers are only right because they *make* the predictions come true. And Sokka finds that extremely irritating.

Consider an event that occurs in the middle of the episode. Sokka approaches a villager who is wearing red shoes. The villager says that Aunt Wu made the following prediction.

(1) When the villager meets their true love, they will be wearing red shoes.

Sokka asks how often the villager wears red shoes. “Every day,” they respond. Sokka becomes furious: “Then of course [the prediction is] gonna come true!” The villager, seemingly unaware of Sokka’s anger, becomes elated: “Really? You think so? I'm so excited!”

At the end of this exchange, Sokka thinks that the villager’s behavior is irrational. To understand why, note that in all likelihood, the villager will fall in love with someone at some point or other. So if they are always wearing red shoes, then (1) is true. But the truth of (1) has nothing to do with any special connection between wearing red shoes and

meeting true loves. If the villager wears red shoes all the time, then for any action X which they will perform, the following holds.

(2) When the villager does X, they will be wearing red shoes.

For instance, the villager will be wearing red shoes when they *fail* to meet their true love. So it is irrational, Sokka thinks, for the villager to constantly wear red shoes: doing so has no impact whatsoever on whether the villager will find true love or not.

As we will see, contrary to Sokka's claims, the villager's behavior is rational. Given that the villager believes (1), it follows from a standard philosophical theory of rational decision-making that in order to be rational, the villager should always wear red shoes. Let us see why.

Expected Utility Theory

The standard philosophical theory of rational decision-making is called "expected utility theory." This theory states the conditions under which someone is rationally required to take a particular action. In other words, expected utility theory says what is—and is not—rational for people to do.

Expected utility theory is based on a particular view of the actions that people perform: for any given person and any given action, there is a certain amount of value—called the "expected value"—which that person can expect to get from performing that action. In other words, the expected value of doing something is what you can reasonably expect to gain or lose by doing that something.

For example, suppose that Katara is playing poker. She is trying to decide between the following two options. First, she could call a bet, and so either gain or lose money depending on who has the better hand. Second, she could fold, and lose all the money she has bet so far. The expected value of calling the bet is the amount of money that Katara can expect to win—or maybe lose—by calling. The expected value of folding is the amount of money that Katara can expect to lose by folding.

There is a formal, mathematical definition of expected value. Because it is fairly complicated, it is presented in an appendix for this chapter. Roughly put, according to that definition, the expected value of an action is a special kind of sum. In particular, the expected value of an action equals a weighted sum of the values of each possible outcome of that action. For example, recall Katara's poker game. The expected value of calling, for Katara, is the sum of two amounts of money. The first amount is how much she expects to get, if she calls and wins. The second amount is how much she expects to lose, if she calls and loses.

Suppose that a person faces exactly two choices: perform an action, or do not perform that action. Then expected utility theory says the following:

> Expected Utility Theory
>
> The person is rationally required to perform the action if and only if the expected value of performing the action is greater, for the person, than the expected value of not performing the action.

In other words, the person should perform that action whenever they expect to get more value from doing so than from not.

Once again, consider Katara playing poker. According to expected utility theory, Katara is rationally required to call the bet if and only if the expected value of calling is greater than the expected value of folding. To put it another way: Katara should call the bet if and only if she can expect to get more money by calling than by folding her hand.

That seems like the right result for a theory of rationality—like expected utility theory—to deliver. For intuitively, if a person expects to get more money by calling, then that is what they should do. So expected utility theory delivers the correct verdict regarding Katara's poker-playing. And that is not a coincidence. Expected utility theory is extremely popular in philosophy because it often gives the right results. That is, expected utility theory often agrees with our intuitions about what agents are rationally required to do.[1]

The Rationality of Wearing Red Shoes

Expected utility theory can show that the villager, discussed earlier, is rationally required to wear red shoes. It follows that Sokka is wrong to suggest that the villager is irrational. For given expected utility theory, the villager is doing exactly what they should.

Recall that according to expected utility theory, an agent is rationally required to perform a particular action whenever the expected value of performing that action is greater than the expected value of not performing that action. So according to expected utility theory, the villager is rationally required to wear red shoes whenever the expected value of wearing red shoes is, for that agent, greater than the expected value of not

wearing them. And so, to determine whether or not the villager's behavior is rational, it suffices to determine whether one of these expected values is greater than the other.

The appendix contains a formal proof which shows that the expected value of wearing red shoes is greater, for the villager, than the expected value of not wearing red shoes. The proof is somewhat complicated, and relies on some mathematical theorems. A summary of the basic ideas underlying the proof will suffice.

The proof relies on just two pieces of information. First, for the villager, we can assume that the value of meeting their true love is greater than the value of not meeting their true love. In other words, the villager values finding love more than not finding it. Second, because the villager believes in Aunt Wu's predictions, they are completely confident that (1)—reproduced below—is true.

(1) When the villager meets their true love, they will be wearing red shoes.

In other words, the villager is completely confident that if they do not wear red shoes, then they will not meet their true love.

With just these two pieces of information, we can show that according to expected utility theory, the villager should always wear red shoes. Very roughly, here is why. Because the villager is so confident in (1), it can be shown that the expected value of not wearing red shoes equals the value, for the villager, of not meeting their true love. That, along with some mathematics, implies the following: the expected value of wearing red

shoes is greater than the expected value of not doing so. Therefore, according to expected utility theory, the villager should always wear red shoes.

In other words, according to expected utility theory, wearing red shoes is rational. Sokka is wrong to criticize the villager for behaving in that way. Given the villager's confidence in Aunt Wu's predictions, wearing red shoes is the right thing to do.[2]

Generalizing to Other Cases

This theory of rationality illuminates other situations in "The Fortuneteller." Time and again, the villagers act in accord with Aunt Wu's predictions. Initially at least, those actions appear to be irrational; they appear that way to Sokka, for instance. But those actions are perfectly rational for the villagers to perform, according to expected utility theory.

Take the very first villager that Aang, Katara, and Sokka encounter, the one being attacked by a platypus bear. Earlier, Aunt Wu predicted that this villager's journey would be safe. Because of their trust in Aunt Wu, the villager is completely certain that this prediction will come true. So, despite the platypus bear's ferocious attacks, the villager remains unfazed. They happily dodge the platypus bear's sharp claws until Appa arrives, roars, and scares the attacker off.

The villager and Sokka have a brief dispute over whether or not it was rational for the villager to remain calm during the attack. Sokka, unsurprisingly, suggests not, for Sokka thinks that Aunt Wu's prediction was wrong: "You didn't have a safe journey; you were almost killed!" Clearly, Sokka thinks that remaining calm was irrational: the villager's calm behavior, Sokka suggests, was motivated by Aunt Wu's false prediction. But as the

villager points out, Aunt Wu's prediction was not false at all: it actually came true. So their journey was safe. Remaining calm, the villager implicitly suggests, was the rational thing to do.

According to expected utility theory, the villager—and not Sokka—is right. Given that the villager is completely confident in the truth of Aunt Wu's prediction, it was rational to remain calm throughout the attack: the expected value of remaining calm is, plausibly, greater than the expected value of panicking.

Wrapping Up

We can draw a general conclusion about the nature of rationality from all this. Sokka assumes that only empirical approaches to reasoning are rational. But given a standard theory of rational decision-making—namely, expected utility theory—that is not so. Mystical approaches to reasoning, such as approaches based on Aunt Wu's fortunetelling, can be rational too.[3]

Because of this, I see "The Fortuneteller" as advancing an argument for the claim that there is no great difference between empirical kinds of reasoning and more mystical kinds of reasoning. At the very least, these two kinds of reasoning do not always differ with respect to how rational they are. Both kinds of reasoning can be rational; and both kinds of reasoning can be irrational too. So rationality is, by its nature, compatible with both reasoning based on empirical considerations and reasoning based on mysticism.

Appendix

In this appendix, I present the fully formal definition of expected value. Then I use that definition to formulate a mathematically precise version of expected utility theory.

Finally, I use all this to prove that the expected value of wearing red shoes is, for the villager, greater than the expected value of not wearing red shoes.

Expected value is defined using propositions and functions. Propositions are represented by lower-case letters: 'x', 'y', and so on. In addition, some propositions are represented using a combination of letters and the negation symbol '¬'. Expressions like '¬x' should be understood as shorthand for "It is not the case that x."

Two different types of functions will be relevant here. One type is used to regiment claims about agents' credences—that is, how confident agents are—in various propositions, given that various other propositions occur. These are called 'credence functions': a credence function is a function Cr which maps pairs of propositions to real numbers between 0 and 1. So if x and y are propositions, then Cr(y|x) is a real number in that numerical range. And the expression 'Cr(y|x)' should be understood as saying "The agent in question has credence Cr(y|x) in the following: given that y occurs, x occurs too."

For example, recall Katara's poker game. Let x be the proposition that Katara wins. Let y be the proposition that Katara calls the bet. And suppose that Cr(y|x)=0.5. Then 'Cr(y|x)=0.5' says the following: Katara has degree of confidence 0.5—in other words, Katara is 50% sure—that if she calls, then she will win.[4]

The other type of function is used to regiment claims about what agents value. These are called 'valuation functions': a valuation function is a function which maps propositions to real numbers. So if y is a proposition, then V(y) is some real number or

other. And the expression 'V(y)' should be understood as saying "The agent in question places V(y) amount of value in the occurrence of y."

For example, let y be the proposition that Katara wins. And suppose that V(y)=2. Then this equation says the following: $2 is how much Katara stands to win.

Now to define expected value. Let x be a proposition which describes an action that an agent might perform. Let y be a proposition which describes a possible outcome of that action. Let Cr be the agent's credence function, and let V be the agent's valuation function. Then the 'expected value' of x—written 'EV(x)'—is defined below.

$$EV(x) = Cr(y|x)V(y) + Cr(\neg y|x)V(\neg y)$$

Roughly put, this equation says that to determine the value which an agent can expect to get from performing the action described by x, you must do the following. First, take the agent's credence that the possible outcome described by y will indeed occur, given that the agent does the action described by x. Multiply that credence by how much the agent values that outcome. The result is, roughly put, the expected value of the specific outcome which 'y' describes. Second, take the agent's credence that the outcome described by y will *not* occur, given that the agent does the action described by x. Multiply that credence by how much the agent values the *non*-occurrence of that outcome. The result is, roughly put, the expected value of the specific outcome which '¬y' describes. Third, add those two numbers together. According to the equation above,

that sum is the overall value which the agent can expect to get, if they perform the action described by x.

Now for the mathematically precise version of expected utility theory. Suppose that an agent faces exactly two choices: perform an action, or do not perform that action. Let x be the proposition that the agent performs that action; so '¬x' expresses the proposition that the agent does not perform that action. Then, according to this version of expected utility theory, the agent is rationally required to perform the action if and only if

$$EV(x) > EV(\neg x)$$

In other words, the agent should perform the action if and only if the expected value of doing so is greater than the expected value of not doing so.

Now let us prove that according to this version of expected utility theory, the villager is rationally required to wear red shoes. Let r be the proposition that the villager wears red shoes, and let m be the proposition that the villager meets their true love. Let Cr be the villager's credence function, and let V be the villager's valuation function.

According to the mathematically precise version of expected utility theory, the villager should wear red shoes if and only if EV(r)>EV(¬r). So to determine whether the villager should wear red shoes, we need only determine whether EV(r) is greater than EV(¬r) or not.

In the main text—in the section called "The Rationality of Wearing Red Shoes"—I mentioned that in order to determine whether or not EV(r)>EV(¬r), just two pieces of information are needed. The first piece of information is expressed by the formula below.

$$V(m) > V(\neg m)$$

In other words, the villager values meeting their true love more than not meeting their true love. The second piece of information is expressed by the following equation.

$$Cr(\neg m|\neg r) = 1$$

In other words, the villager is completely confident that if they do not wear red shoes, then they will not meet their true love.[5] Both pieces of information are, in fact, true: the villager really does value meeting their true love more than not, and the villager really is completely confident that if they do not wear red shoes, then that meeting will not occur. So in the proofs to come, it is reasonable to assume both that V(m)>V(¬m) and that Cr(¬m|¬r)=1.

Now to prove that EV(r)>EV(¬r). A standard theorem of probability, in conjunction with the fact that Cr(¬m|¬r)=1, implies that Cr(m|¬r)=0.[6] Therefore,

$$\begin{aligned} EV(\neg r) &= Cr(m|\neg r)V(m) + Cr(\neg m|\neg r)V(\neg m) \\ &= 0 \cdot V(m) + 1 \cdot V(\neg m) \end{aligned}$$

$$= V(\neg m)$$

It follows that EV(r)>EV(¬r) if and only if EV(r)>V(¬m). So to complete the proof, it suffices to show that 'EV(r)>V(¬m)' holds.

Towards that end, multiply both sides of the inequality 'V(m)>V(¬m)' by 'Cr(m|r)'. The result is below.[7]

$$Cr(m|r)V(m) > Cr(m|r)V(\neg m)$$

By the theorem of probability mentioned earlier, Cr(m|r)+Cr(¬m|r)=1; in other words, Cr(m|r)=1- Cr(¬m|r). Substituting '1- Cr(¬m|r)' for 'Cr(m|r)' in the right side of the above inequality yields

$$Cr(m|r)V(m) > \big(1 - Cr(\neg m|r)\big)V(\neg m)$$

By some straightforward algebra, it follows that

$$Cr(m|r)V(m) + Cr(\neg m|r)V(\neg m) > V(\neg m)$$

And since EV(r)=Cr(m|r)V(m)+Cr(¬m|r)V(¬m), it follows that EV(r)>V(¬m). This completes the proof.

[1] For a thorough introduction to expected utility theory, and philosophical theories of decision-making more generally, see Martin Peterson, *An Introduction to Decision Theory* (New York: Cambridge University Press, 2017) and Katie Steele and H. Orri Stefánsson, "Decision Theory," *The Stanford Encyclopedia of Philosophy*, Winter 2020, at https://plato.stanford.edu/entries/decision-theory/.

[2] It is worth pointing out that expected utility theory has a striking implication: if a person happens to be extremely confident in a liar, or a quack, or a conspiracy theorist, then it is often rational for that person to believe whatever that liar, or quack, or conspiracy theorist, says. Though this is clearly a problematic implication, philosophers often disagree over what exactly the problem is. Some philosophers take this to be a theoretical problem. It shows that the theory of expected utility is wrong, because the correct theory of rationality would not have this implication. Other philosophers take this to be a practical problem. Rather than showing that expected utility theory is wrong, it shows that in order to avoid being misled, people need to be taught more than just how to think rationally: the practical problem is the problem of determining what else to teach people, and how to teach it.

[3] One might claim that the villages' reasoning should actually count as empirical. The reasoning itself is not mystical, one might claim: rather, mysticism only enters the villagers' reasoning insofar as the source of their information—namely, Aunt Wu—is mystical (thanks to William Irwin for suggesting this). This approach to the villagers' reasoning is compatible with my own. For this approach still draws a distinction between two kinds of reasoning: namely, reasoning on the basis of empirical sources only, and reasoning on the basis of at least some mystical sources. Understood in this way, "The Fortuneteller" can be interpreted as advancing an argument for the claim that there is no great difference between these two kinds of reasoning. Rationality, in other words, is compatible with reasoning based on mystical sources.

[4] Functions like these are often called 'conditional credence functions.' For more details about credence functions, and other notions from probability theory formulated over spaces of propositions expressed using logical symbols like '¬', see E. T. Jaynes, *Probability Theory* (New York, NY: Cambridge University Press).

[5] This equation is a formal representation of the fact that the villager is completely confident in the truth of (1) from the section called "Sokka, the Villagers, and Aunt Wu" in the main text. That is, the villager is completely confident in the truth of Aunt Wu's prediction.

[6] It is usually assumed that credence functions like Cr obey the axioms of probability theory. That is why a probability theorem can be used here; and that is why facts about probabilities are used elsewhere in this proof. For an explanation of why credence functions like Cr are often assumed to obey the probability axioms, see Alan Hájek, "Dutch Book Arguments," in Paul Anand, Prasanta K. Pattanaik, and Clemens Puppe (Eds.), *The Handbook of Rational and Social Choice* (New York, NY: Oxford University Press, 2009), 173–195.

[7] To obtain this result, I assumed that $Cr(m|r)>0$. This assumption is definitely reasonable for the villager in question. They definitely have nonzero credence that if they wear red shoes then they will meet their true love.