

1-1-2003

# Being and Acting Rational

Michael N. Huhns

University of South Carolina - Columbia, huhns@sc.edu

Follow this and additional works at: [http://scholarcommons.sc.edu/csce\\_facpub](http://scholarcommons.sc.edu/csce_facpub)



Part of the [Computer Engineering Commons](#)

---

## Publication Info

Published in *IEEE Internet Computing*, Volume 7, Issue 2, 2003, pages 91-93.

<http://ieeexplore.ieee.org/servlet/opac?punumber=4236>

© 2003 by the Institute of Electrical and Electronics Engineers (IEEE)

This Article is brought to you for free and open access by the Computer Science and Engineering, Department of at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact [SCHOLARC@mailbox.sc.edu](mailto:SCHOLARC@mailbox.sc.edu).

# Being and Acting Rational



Michael N. Huhns • University of South Carolina • huhns@sc.edu

I usually prefer to deal with rational people. In the vernacular, being *rational* means being reasonable – that is, using reason when making decisions, taking actions, and achieving goals. Rational people tend to be more predictable and, therefore, understandable. Understanding a decision makes it easier for me to accept it.

When I encounter and interact with agents on the Web, or when my own agents encounter them, I prefer that these agents behave rationally as well. An agent can behave rationally in at least three ways. Some of these might be better than others; some might be appropriate for individuals but not for groups.

## Rationality Types

A rationality theory indicates what is rational and what is not in specific cases. Three such theories that govern an agent's behavior are *logical rationality*, *economic rationality*, and *pragmatic rationality*. They depend, respectively, on the mathematics of logic, probability, and computation. To conform to their corresponding mathematical formalism, each type requires strong assumptions about the nature of a rational agent's world and how the agent can sense and act in that world.

### Logical Rationality

In the fourth century BC, the Greek philosopher Aristotle formulated a process for reasoning that would lead to irrefutable conclusions. His system of *syllogisms* (a kind of inference mechanism) could produce not only new knowledge, but also a means to achieve goals based on logically justified actions.

The resultant view – that reasoning could be specified precisely and thus mechanized – evolved in the early 20th century into the doctrine of *logical positivism*, which held that everything an agent knows can be derived from observation sentences that represent the agent's environment. The derivations come from applying laws of deductive logic.

Deductive logic provides rational constraints on

belief in two ways. First, it can be used to define the notion of deductive consistency and inconsistency: deductive inconsistency determines a kind of incoherence in belief. Second, the laws of deductive logic can constrain admissible changes in belief by providing *deductive rules of inference*. For example, *modus ponens* is a deductive rule of inference that requires  $Q$  to be inferred from sentences  $P$  and  $P \rightarrow Q$ .

So, to be logically rational, an agent “simply” has to convert everything it senses into a sentence (a belief) in a formal language, combine the sentences with all other sentences it has sensed or derived, derive new sentences about its world, and use this new set of sentences to choose its actions.

Several major problems cloud this approach. First, observations about the world might be uncertain and incomplete, making them difficult to express logically. Second, several courses of action could lead to a goal's achievement, and it is difficult for logic to help an agent decide among them. Third, there might not be any action that an agent can prove will achieve its goal, leaving the agent without help in deciding what to do. Finally, reasoning about a large set of sentences might be intractable.

### Economic Rationality

Another option is for agents to be economically rational.<sup>1</sup> Like logically rational agents, economically rational ones act to achieve their goals on the basis of what they know. Operationally, however, an economically rational agent ranks possible actions by the expected utility of their results and then executes the action that has the highest expected utility. (Expected utility is defined in terms of the agent's possible actions, the probabilities of the actions' outcomes, and the agent's ranked preferences among those outcomes.) Put simply, economic rationality is based on decision theory, which combines logic and probability theory with utility theory to provide a means for making decisions under uncertainty.

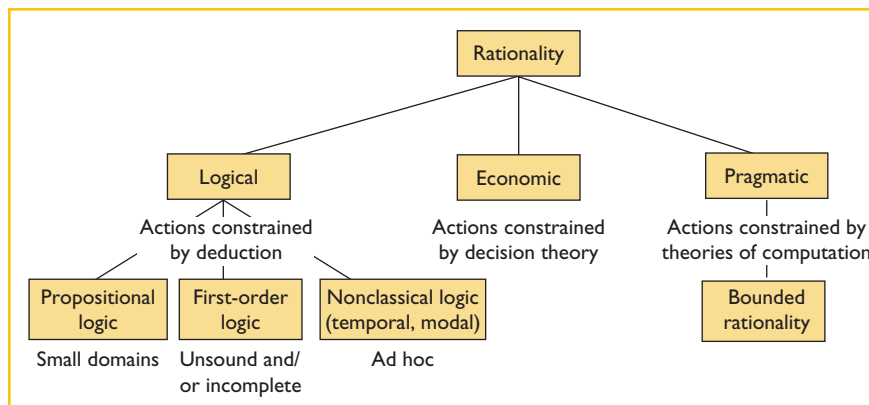


Figure 1. Rationality types and their major characteristics. Formal approaches to rationality, such as logical and economic rationality, have severe limitations when scaled up and applied in real-world domains. Theories of bounded rationality are more closely related to human reasoning and appear more promising for practical implementations of agents.

Applying probability theory to rationality is attributed to the Reverend Thomas Bayes (c. 1701–61).<sup>2</sup> Bayesian epistemology’s two main features are the introduction of a *formal apparatus* for inductive logic and the introduction of a *pragmatic self-defeat test* for epistemic rationality as a way to extend justification of the laws of deductive logic to include inductive logic.<sup>3</sup> The formal apparatus adds standards of probabilistic coherence and a rule of probabilistic inference, both of which apply to degrees of belief (degrees of confidence). Bayesian decision theory is now the dominant theoretical model for both the descriptive and normative analyses of decisions.

Unfortunately, economic rationality requires a computationally expensive search over the outcomes of all possible sequences of actions (because several actions might be needed to achieve a goal), knowledge of the probability distributions for the outcomes (which are difficult to determine), and a means for assigning utilities to outcomes.<sup>4</sup>

**Pragmatic Rationality**

The earlier approaches to rationality rely on the assumption that the world will not significantly change while the agent decides what to do, and that an action that is rational when decision-making begins will be rational

when it concludes. Clearly, this is problematic in real-world settings. Imagine a car-driving agent waiting at a stop sign. The agent looks both ways, does not see any other vehicles, and then, remembering that it dropped its map, rummages on the floor for several minutes to find it. Finally, with map in hand, it deduces from its observations that there are no approaching vehicles and drives blithely across the intersection.

Similarly, if proving that it was safe to cross the intersection took longer than the time for the traffic status to change, the resultant proof would be worthless. A pragmatic approach takes such computational limitations into account.

According to Stuart Russell and Peter Norvig,<sup>2</sup> this means doing “the right thing.” Formally, they define this as “For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.”

When agents have limited abilities, the best they can achieve is bounded rationality. The ideal of logical or economic rationality requires choosing actions to maximize a measure of expected utility. That utility should reflect a complete and consistent preference order and probability measure

over all possible contingencies. This requirement appears too strong to permit an accurate specification for realistic individual agent behavior.

We can weaken the ideal requirements for rationality in many ways. Possibilities include anytime algorithms that return the best action found each time they terminate and theories that attempt to mimic human decision making. Because of the rich variety of psychological types we can observe in humans – each with different reasoning strengths and limitations – it is unlikely that there will be a single best approach to pragmatic rationality.<sup>5</sup>

**Rationality and Multiagents**

Rationality is important for groups of agents as well. If all agents in a group are individually rational, no matter what type of rationality they use, will the group necessarily behave rationally? That is, will the group always make the decisions, take the actions, and achieve the goals that are best for it? This question is important for governments, political organizations, corporations, teams, and committees.

An agent’s best strategy often depends on what strategies other agents choose. For each agent in a group to behave rationally by maximizing its self-interest, for example, it must consider the behaviors of other agents who are also behaving in their own self-interest. This consideration is the basis of game theory, which provides mathematical guidance for how agents in a multiagent system decide their actions.

In open or continuous environments, deciding what is best depends on a time horizon – it is usually impractical for agents to reason infinitely far into the future or to consider an infinite number of intermediate states. For a given finite time horizon, an agent must choose a strategy that considers either the consequences of just the end result or the consequences of both the ends and the means. When the agents are part of a society, ethics can provide some guidance.

The ethical theory *egoism* holds that action should maximize self-interest. A parallel theory called *utilitarianism* holds that action should maximize the universal good of all agents. Both theories consider only the end result (they are *teleological*); they hold that the best thing to do is always maximize a certain good, in which good can be interpreted as pleasure, preference satisfaction, interest satisfaction, or aesthetic ideals. In contrast, *deontological* theories hold that the ends do not justify the means, and agents must at each step choose the action that does not endanger society's welfare.

### Conclusion

Rationality alone is insufficient to specify agent design. Using economic theory, we can program agents to behave in ways that maximize their utility while responding to environmental changes. However, economic models for agents, although general in principle, are typically limited in practice because the value functions that are tractable essentially reduce an agent to acting selfishly.<sup>2</sup> Building a stable social system from a collection of agents motivated by self-serving interests is difficult.

Finally, understanding rationality and knowledge requires interdisciplinary results from artificial intelligence, distributed computing, economics and game theory, linguistics, philosophy, and psychology. A complete theory involves semantic models for knowledge, belief, action, and uncertainty; bounded rationality and resource-bounded reasoning; commonsense epistemic reasoning; reasoning about mental states; belief revision; and interactions in multiagent systems. □

### Acknowledgment

The US National Science Foundation supported this work under grant number IIS-0083362.

### References

1. M.N. Huhns and L.M. Stephens, "Multiagent Systems and Societies of Agents," *Multiagent Systems*, G. Weiss, ed., MIT Press, 1999, pp. 106–111.
2. S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 2nd ed., Pearson Education, 2003.
3. W. Talbott, "Bayesian Epistemology," *The Stanford Encyclopedia of Philosophy*; <http://plato.stanford.edu/archives/fall2001/entries/epistemology-bayesian>.
4. M. Wooldridge, *Reasoning About Rational Agents*, MIT Press, 2000.
5. J. Doyle, "Rational Decision Making," *MIT Encyclopedia of Cognitive Science*; <http://cognet.mit.edu/MITECS/Articles/doyle2.html>.

**Michael N. Huhns** is a professor of computer science and engineering at the University of South Carolina, where he also directs the Center for Information Technology.

Advertiser / Product	Page Number
John Wiley & Sons	Back Cover

### Advertising Personnel

<b>Marion Delaney</b> IEEE Media, Advertising Director Phone: +1 212 419 7766 Fax: +1 212 419 7589 Email: md.ieeemedia@ieee.org	<b>Sandy Brown</b> IEEE Computer Society, Business Development Manager Phone: +1 714 821 8380 Fax: +1 714 821 4010 Email: sb.ieeemedia@ieee.org
<b>Marian Anderson</b> Advertising Coordinator Phone: +1 714 821 8380 Fax: +1 714 821 4010 Email: manderson@computer.org	<b>Debbie Sims</b> Assistant Advertising Coordinator Phone: +1 714 821 8380 Fax: +1 714 821 4010 Email: dsims@computer.org

### Advertising Sales Representatives

<b>Mid Atlantic (product/recruitment)</b> Dawn Becker Phone: +1 732 772 0160 Fax: +1 732 772 0161 Email: db.ieeemedia@ieee.org	<b>Southeast (product/recruitment)</b> C. William Bentz III Email: bb.ieeemedia@ieee.org Gregory Maddock Email: gm.ieeemedia@ieee.org Sarah K. Wiley Email: sh.ieeemedia@ieee.org Phone: +1 404 256 3800 Fax: +1 404 255 7942
<b>Midwest (product)</b> David Kovacs Phone: +1 847 705 6867 Fax: +1 847 705 6878 Email: dk.ieeemedia@ieee.org	<b>Midwest/Southwest recruitment)</b> Tom Wilcoxon Phone: +1 847 498 4520 Fax: +1 847 498 5911 Email: tw.ieeemedia@ieee.org
<b>New England (product)</b> Jody Estabrook Phone: +1 978 244 0192 Fax: +1 978 244 0103 Email: je.ieeemedia@ieee.org	<b>New England (recruitment)</b> Barbara Lynch Phone: +1 401 739 7798 Fax: +1 401 739 7970 Email: bl.ieeemedia@ieee.org
<b>Southwest (product)</b> Royce House Phone: +1 713 668 1007 Fax: +1 713 668 1176 Email: rh.ieeemedia@ieee.org	<b>Connecticut (product)</b> Stan Greenfield Phone: +1 203 938 2418 Fax: +1 203 938 3211 Email: greenco@optonline.net
<b>Northwest (product)</b> John Gibbs Phone: +1 415 929 7619 Fax: +1 415 577 5198 Email: jg.ieeemedia@ieee.org	<b>Southern CA (recruitment)</b> Tim Matteson Phone: +1 310 836 4064 Fax: +1 310 836 4067 Email: tm.ieeemedia@ieee.org
<b>Southern CA (product)</b> Marshall Rubin Phone: +1 818 888 2407 Fax: +1 818 888 4907 Email: mr.ieeemedia@ieee.org	<b>Japan</b> German Tajiri Phone: +81 42 501 9551 Fax: +81 42 501 9552 Email: gt.ieeemedia@ieee.org
<b>Midwest (product)</b> Dave Jones Phone: +1 708 442 5633 Fax: +1 708 442 7620 Email: dj.ieeemedia@ieee.org Will Hamilton Phone: +1 269 381 2156 Fax: +1 269 381 2556 Email: wh.ieeemedia@ieee.org	<b>Europe (recruitment)</b> Penny Lee Phone: +44 020 7405 7577 Fax: +44 020 7405 7506 Email: PennyLee@Essentialmedia.Co.Uk
Joe DiNardo Phone: +1 440 248 2456 Fax: +1 440 248 2594 Email: jd.ieeemedia@ieee.org	<b>Europe (product)</b> Hilary Turnbull Phone: +44 131 660 6605 Fax: +44 131 660 6989 Email: impress@impressmedia.com