24.223 Rationality

Fall 2023
Lecture: M/W 11–12:30
Classroom: 56-162
Course website: kevindorst.com/fa23_rationality

Professor: Kevin Dorst
kmdorst@mit.edu
Office: 32-D931
Office Hours: M 1–2

Class Description: How rational are people? Some bits of cognitive science suggests we're surprisingly foolish; other bits suggests we're amazingly smart. What should we make of these competing pictures?

Answering this question requires combining both normative and descriptive methods: we need to know how rational people *would* reason, as well as how real people *do* reason. This course will begin with the normative, and then bring it to the descriptive. We'll spend the first part of the course learning the basics of Bayesian theories of rationality. We'll then use those normative tools to help think through the proper interpretation of a variety of empirical results that have been taken to demonstrate human irrationality. For example: how rational is hindsight bias, sunk-cost reasoning, the gambler's fallacy, conformity, polarization, and so on?

Course Goals: Addressing these questions involves combining mathematical reasoning with philosophical reflection. There's an art to doing this well. We'll try to learn it.

Readings:

Textbook. You will need **both** the **first** and **second** volumes of Titelbaum, *Fundamentals of Bayesian Epistemology*. The hardback is expensive; buy the paperback.

Papers. All other readings are posted in a zip file on the course website: kevindorst.com/fa23_rationality

Grading: Your grade will be determined by attendance and participation (10%), 2 problem sets (20%), and two papers (70%).

Attendance/participation (10%). Attendance is mandatory. This will be a partially discussion-based class, so to receive full participation credit you must also contribute at least semi-regularly. If you have trouble speaking up in class, you can instead email me questions before class.

Problem sets (2 psets, 10% each). Submitted in class. You will be given problem sets 1 weeks before they are due. They will be drawn from the Titelbaum textbook, and will consist in both mathematical and conceptual questions. No prior courses are required (though logic—and of course probability—will make it easier).
Working with others is a good idea, but you must write up your answers on your own.
This course uses **Pset Partners**: https://psetpartners.mit.edu/

Papers. Submitted via email to kmdorst@mit.edu. Writing philosophy papers is a learned skill that can be tough to get the hang of. It requires making a single contribution to an ongoing conversation started by the readings and in class. We'll devote class time to discussing what's expected, and I'll send out materials beforehand. Prompts will be distributed.

First paper (30%). A roughly 1500-word (5–6 page) paper in response to one of the readings.
Second paper (40%). A roughly 2000-word (7–8 page) paper in response to one of the readings.

IMPORTANT: ChatGPT and AI Policies.
You are free to use ChatGPT or other AI models in the development and brainstorming of your ideas—just as you would talk through your paper ideas with a friend, or use a search engine to learn about a topic. **However, you must write and edit your paper yourself. Using ChatGPT to write (parts of) your paper for you will count as plagiarism.**
The reasoning: You may never again need to come up with your own arguments or ideas *purely* from scratch, without AI assistance. You certainly **will** need to learn how to develop, formulate, frame, and express your own ideas and arguments in your own words. That is the skill that this course is trying to teach.
The method: Draftback. To help maintain your willpower in the midst of the semester, **you are required to write your entire paper—from start to finish—in Google Docs.** (You can outline and brainstorm in whatever application you like.) When you submit the final version, you will also **share a Google Doc link, giving me editing privileges**. This will allow me to use the "Draftback" Chrome plugin to verify that the paper was written by a human—it creates a time lapse of the keystroke-by-keystroke writing of the paper, allowing me to quickly verify the that it involved all the typos, false starts, and sentence- and paragraph-rewrites characteristic of human writing. I will only check the Draftback time-lapse if I have independent reason to suspect that the paper was written by an AI. No part of your writing process will be evaluated.
Failing to submit a Google-Doc link will provide (Bayesian) evidence that the paper is AI-generated. **If it is AI-generated, you will automatically fail the course and be referred to the academic integrity office.** It's not worth it.

Course Policies:

Announcements. I'll send class communications via email. Make sure to give me an email that you check regularly.

Late work. Late work will be marked down 1/3 of a letter grade for each day late (a B+ becomes a B, etc.). Problem set answer-keys will be provided 1 week after the due date, so **problem-sets turned in after this will receive a zero.**

Writing Help. Clear and concise writing is an important component of this class. The MIT Writing and Communication Center (WCC) offers free one-on-one professional advice from communication specialists with advanced degrees and publishing experience. The WCC can help you learn about all types of academic and professional writing and further develop your oral communication skills. You can learn more about WCC consultations at http://cmsw.mit.edu/writing-and-communication-center and register with the online scheduler to make appointments through https://mit.mywconline.com. WCC hours are Monday-Friday, 9am–6pm during the semester, and fill up fast.

Academic Integrity. Any writing submitted for a grade must be your own; quotations or ideas paraphrased from other sources should be clearly cited in a way that allows us to find the source. You are responsible for knowing what counts as plagiarism. See https://integrity.mit.edu/sites/default/files/images/AcademicIntegrityHandbook2020-color.pdf. Reach out if you have questions. Asking beforehand is never an issue.

Disability Accommodations. MIT is committed to the principle of equal access. Students who need disability accommodations are encouraged to speak with Disability and Access Services (DAS), prior to or early in the semester so that accommodation requests can be evaluated and addressed in a timely fashion. If you have a disability and are not planning to use accommodations, it is still recommended that you meet with DAS staff to familiarize yourself with their services and resources. Please visit the DAS website for contact information: https://studentlife.mit.edu/das.

If you've already been approved for accommodations, please inform me ASAP so we can make an implementation plan.

Schedule

I. Rationality under uncertainty

1. (Sept 6) Gershman 2021, pp 1–6
2. (Sept 11) Degrees of belief, Titelbaum Ch. 1
3. (Sept 13) Probability, Titelbaum Ch. 2
   Problem set 1 distributed.
4. (Sept 18) Conditional Probability, Titelbaum Ch. 3
5. (Sept 20) Conditionalization, Titelbaum Ch. 4
6. (Sept 25) Deference Principles, Titelbaum sections 5.1–5.3
   Problem Set 1 due in class
7. (Sept 27) Confirmation, Titelbaum Ch. 6
   Problem set 2 distributed.
8. (Oct 2) Decision Theory, Titelbaum Ch. 7

II. Human (ir)rationality

9. (Oct 4) Hindsight bias, Hedden 2019
10. (Oct 9) No class (student holiday).
11. (Oct 11) The gambler's fallacy, Dorst blogpost (2020)
    Problem Set 2 due in class
12. (Oct 16) Motivated reasoning, Mandelbaum 2019
    Paper 1 prompts distributed
13. (Oct 18) Motivated Ignorance, Kinney and Bright 2022
14. (Oct 23) The sunk cost fallacy, Kelly 2004
15. (Oct 25) No class. (I'm traveling.)
16. (Oct 30) Overconfidence, Dorst 2023
17. (Nov 1) Polarization, Henderson and Gebharter 2021

18. (Nov 6) Polarization, Nielsen and Stewart 2018, Part 1 (pages 1–10)
    Paper 1 due at 11:59pm. Remember to share permissions to your google doc.
19. (Nov 8) Polarization, O'Connor and Weatherall 2018
20. (Nov 13) Polarization, Kelly 2008
21. (Nov 15) Confirmation bias, Dorst 2023, sections 1–2 **and** Salow 2018, section 1
22. (Nov 20) Confirmation bias. Hahn and Harris 2014
    Paper 2 prompts distributed
23. (Nov 22) No class. (Thanksgiving)
24. (Nov 27) Selective memory, Singer et al 2019
25. (Nov 29) Selective memory, Dallmann 2017
26. (Dec 4) Selective memory, Wilson 2014
27. (Dec 6) Wishful thinking, Melnikoff and Strohminger 2023
28. (Dec 11) The explore/exploit tradeoff, Aronowitz 2021
29. (Dec 13) Catch-up / review
    Paper 2 due at 11:59pm. Remember to share permissions to your google doc.