

The Bounded Rational Frontier: Implications for research in AI, Economics and Philosophy

Jonathan Harris

September 22, 2023

jonathan@total-portfolio.org

This note introduces a toy model of bounded-rational decision making that demonstrates how open-mindedness and uncertainty aversion can be rational under practical constraints. This has important implications for research in a range of fields, from the design of AI systems and policies to the management of philanthropic and impact investing portfolios. The results are visualized as 'frontier' curves, providing insights into the complexities of bounded rationality.

Introduction

Consider the following decision problem: You have \$1,000 to give to charity. Conventional thinking would recommend giving the whole amount to the 'best' charity. That keeps things simple and is optimal if you can identify the best charity. Yet, you find yourself hesitating. You feel almost clueless about how to compare different impacts – from improving global health, to reducing animal suffering, to mitigating catastrophic risks from AI. You encounter a wide range of conflicting views while doing your research. And you realize that you can give to multiple charities. You've even noticed that your friends who care a lot about donating effectively often give to several charities. Is that a reasonable thing to do in the face of uncertainty?

This scenario illustrates a broader challenge: how can we ensure that decision-makers, whether human or artificial, make optimal choices when their resources—time, expertise, and information—are constrained? While classical models often assume perfect rationality, real-world decision-makers always operate under significant constraints. They can't be rational in an idealistic sense, but they can be 'bounded rational' (e.g., Genewein et al., 2015) or 'resource rational' (e.g., Bhui, Lai, Gershman (2021), Lieder, Griffiths (2019)) - that is, following the optimal policy given the practical constraints that they face. For simplicity, I will use the term 'bounded rational' to refer to all related frameworks.

This note presents a simple toy model of bounded-rational decision making to highlight three key concepts that are often overlooked:

1. **Decision making as inference.** The decision maker combines multiple imperfect models to *infer* the true state of the system, producing better results than relying on a single model.
2. **Rational Uncertainty Aversion.** The more volatile the potential outcomes, the more the decision maker will tend to underperform what they would achieve in ideal conditions. This can make it rational for them to prefer less risky, less ambiguous options and to diversify.
3. **Frontier Curves.** The results are visualized as curves that show how different strategies trade off expected outcomes against risk and other factors—similar to the efficient frontier used in finance. These curves are a powerful tool to make the results of different models more transparent.

Given the rapid advances in AI and the increasing complexity of global challenges, addressing bounded rationality and its implications is more urgent than ever. These concepts are important because:

- The mental models we walk around with can influence everything we do from how we engage with others about policy to how we design technical AI systems. 'Inference' models where it is best to

combine multiple points of view can serve as powerful reminders of the importance of open-mindedness in these times of increasing polarization and conflict. Yet, despite their intuitive appeal, such models are not the norm. The dominant paradigm is more for different models to compete with each other to be the one that makes the decision, without any real mixing of models occurring.

- The potential for uncertainty aversion to be rational has implications for a wide range of issues, from perennial ones like how much to give to charity to urgent ones like AI alignment. Yet, it has been rarely, if ever, explored.
- Much of the related literature often lacks clear visualizations, like the 'frontier' curves, even though this is important to ensure effective communication to an interdisciplinary audience.

This note is not a final paper but an invitation to researchers from various disciplines to collaborate in exploring these important concepts. The goal of the note is to demonstrate the importance of these concepts in a simplified setting that can be adapted to a range of fields in the future, including philosophy and economics. By doing so, this note aims to spark further exploration and collaboration among researchers in AI, economics, philosophy, and related fields. I am actively seeking collaborators who are interested in building on this work to formalize, extend, or adapt it to their specific areas of expertise.

The structure of this note is as follows. First, the basic inference setup of the toy model is introduced. We then explore how the results of bounded rational inference can be visualized using frontier curves. Next, we see how rational uncertainty aversion can emerge from constraints on the decision makers estimation abilities. Finally, we discuss the implications of these findings for various fields and suggest directions for future research.

Basic Inference Setup

The setup, similar to Buchak (2023), is that a person has committed to donate a total of \$1,000 to two charities:

- Charity A - A benchmark charity that produces 1 unit of impact per dollar.
- Charity B - An alternative charity that produces either $s = 0$ or $s = v > 1$ units of impact per dollar, where v is a known, fixed parameter.

They can choose the fraction a to donate to charity B in increments of one-tenth. For example, if they choose $a = 0.7$ and B is in state $s = 2$, then the impact per dollar of their donation will be $0.3 * 1 + 0.7 * 2 = 1.7$.

Without additional information, the probability of state $s = v$ is such that the expected impact of both charities is equal to 1. This means that higher values of v correspond to lower upside probabilities for charity B, making v a measure of the associated risk level.

To inform their decision, the donor has access to two analysts who provide different observations about charity B's state:

- One provides a quantitative score (0, 1, or 10)
- The other provides a qualitative rating (x, y, or z)

The donor has a probabilistic model for the relationship between observations and states. For each pair of observations, they can estimate the probability of each state and use this to calculate the expected impact of each action across both possible states. This approach allows them to combine observations from the analysts even if they aren't easily comparable (i.e., quantitative and qualitative).

In this note we consider several observation scenarios:

ID	Scenario Name	Description
0	Perfect information	At least one analyst identifies the state with 99% accuracy.
1	Good, independent observations	Both analysts have a 'good' ability to identify the state and are otherwise uncorrelated.
2	Correlated observations	Same as Scenario 1 but the analysts views are partially correlated (e.g., using similar methods and evidence).
3	Only one good observation	One analyst is good, the other is clueless; or equivalently, both analysts are 100% correlated.
4	Relatively clueless	One analyst is clueless, the other offers only a slight edge in identifying the true state.

Figure 1 illustrates the donor's model for the observations given the states for the 'Good, independent observations' scenario. Similar charts for other scenarios can be generated using the code provided at the end of this notebook.

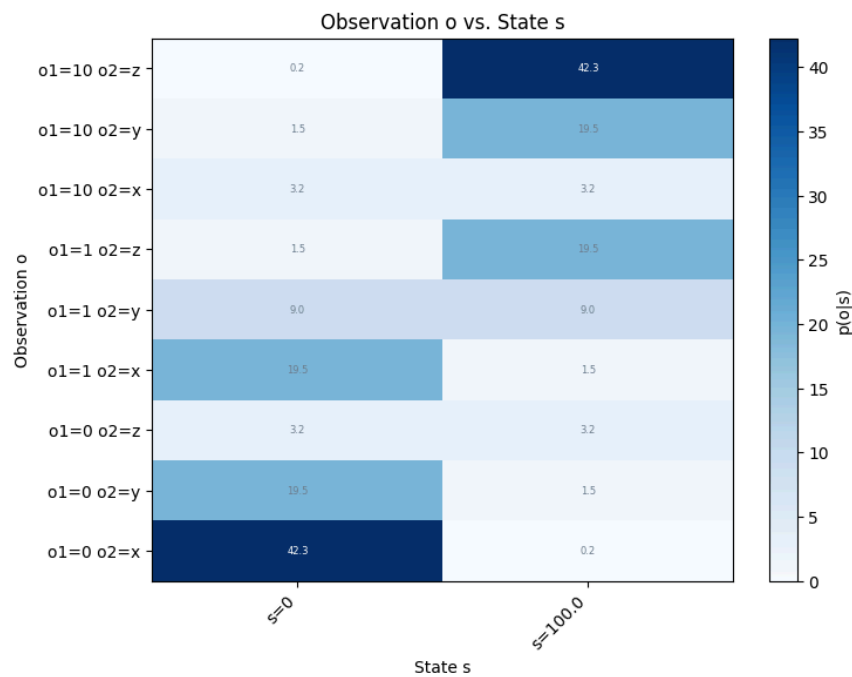


Figure 1: The donor's beliefs about the probabilities of the observations given the states for the "Good, independent observations" scenario (with $v=100$).

Bounded Rational Inference

The donor in our setup needs to choose their action a given the observations o . Generalizing, we can think of their policy as the probability $p(a|o)$ that they take action a given observation o . In our simple, linear toy model, for a given observation o , an ideal rational agent will always have $p(a^*|o) = 1$ for some action a^* and the rest of the probabilities as zero. A key insight of bounded rationality is that such policies are more 'complex' in the sense of require more computations to implement than randomly choosing an action.

Implementing a policy for the toy model may seem trivial, but in general an ideal policy requires more computational 'bandwidth'. It requires storing the policy in memory, capturing the observations, looking up the corresponding action and executing it. Whereas a random strategy just requires randomly choosing an action and executing it.

Thus, one of the most popular and promising models for bounded rationality adds a cost for policy complexity (see, for example, Genewein et al. (2015) and Lai and Gershman (2024)). Figure 2 shows the optimal policy for different levels of complexity costs. It illustrates how higher complexity costs increase the randomness of the optimal policy, thereby decreasing its expected impact. This leads us to explore the concept of frontier curves in the next section, as a way to visualize the trade-off between complexity and expected impact.

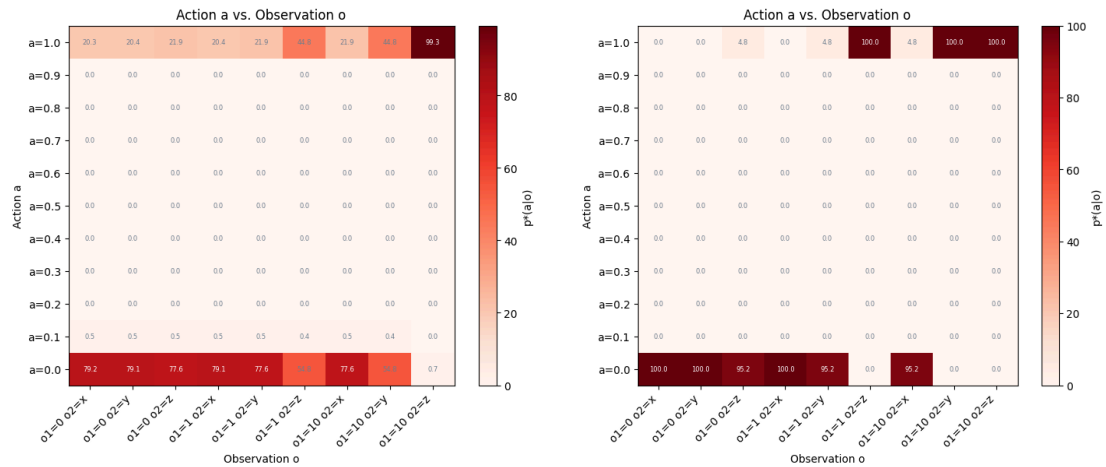


Figure 2: Optimal action probabilities given each observations for the "Good, independent observations" scenario with low complexity/high costs (left) and high complexity/low costs (right). The low complexity policy is much more random. The high complexity policy still includes some randomness because for some observations the expected values of the actions are very similar.

Frontier Curves

A natural way to present the results of information-theoretic bounded rationality models is with complexity-utility frontier curves. These curves show the 'frontier' of the minimum complexity required to achieve a certain level of expected impact or equivalently, the maximum achievable expected impact given a level of complexity. This is similar to the 'Efficient Frontier' volatility-return curve in finance and the 'Efficient Impact Frontier' impact-return curve in impact investing. These visualizations are a powerful way to understand the implications of different forms of bounded rationality.

Figure 3 presents these frontier curves for the toy model. It illustrates that:

- The optimal Expected Impact strictly increases as Policy Complexity is allowed to increase.
- Higher expected impact can be consistently obtained in the scenarios where the donor has good, alternative observations to combine ('Good, independent observations' and 'Correlated observations').
- The value of combining different views together is significantly reduced when those views are not independent ('Correlated observations').

These results show how a basic model of bounded rational inference can work. However, this form of bounded rationality does not generate uncertainty aversion. Rather it just forces less complex policies to be more random, which may or may not result in behavior that looks like uncertainty aversion. So, in the next section we turn the complexity cost all but off and explore a modification to the model that generates uncertainty aversion.

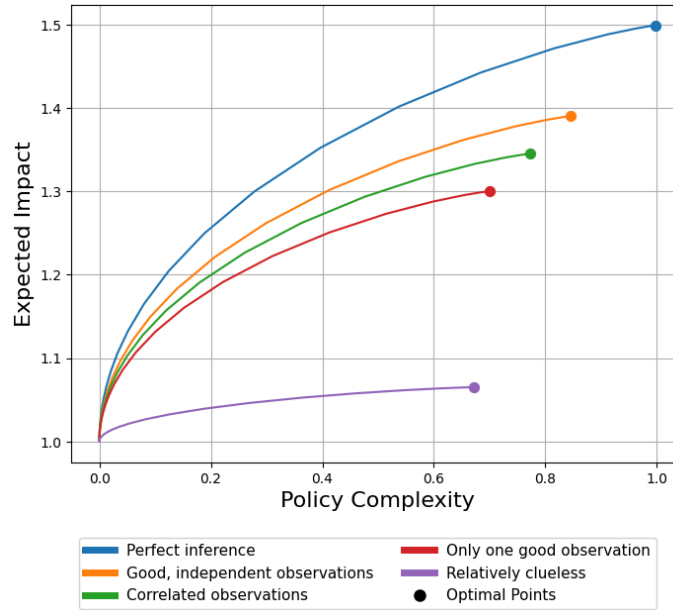


Figure 3: Expected impact-complexity frontier curves for each observation scenario. Expected Impact is the expected impact per dollar donated and Policy Complexity is the '*mutual information*' between the actions and observations for the policy.

Rational Uncertainty Aversion: When Playing It Safe Pays Off

Bounded rationality models have largely focused on limitations to metrics like 'Policy Complexity'. This overlooks a crucial consideration for how the optimal policy is computed in practice: estimation error in the expected value of each action. While this is a practical issue in machine learning, exemplified by systems like DeepMind's Go-playing AI which can only sample from a fraction of the possible games, it has received surprisingly little attention in the broader bounded rationality literature. By incorporating estimation error into bounded rationality models, we can better reflect the realities faced by both artificial and human decision-makers.

We add this issue to the toy model by supposing that the donor can't directly compute the expected impact of each action. Instead, they must estimate it by randomly sampling the possible outcomes and taking an average across the samples as in Lieder, Hsu & Griffiths (2014). This may seem artificial in the context of our toy model, but it would be reality in more complex scenarios with many possible states and actions.

Policies that are optimized based on noisy expected impact estimates will be strictly worse than the true optimal policy. If the performance gap increases significantly in riskier scenarios, then low-risk scenarios could offer higher expected impact in practice even if higher-risk scenarios would offer much higher expected impact under ideal conditions.

To test the results we restrict the donor to 2, 8, 32 or 128 samples for their estimates. The following subsections illustrate how these sampling constraints can generate different forms of uncertainty aversion. For more detail, see the brief technical explanation at the bottom of this note.

Aversion to riskier bets

Figure 4 presents frontier curves that show the expected impact for different values of v up to $v = 100$. With perfect information or a large number of samples the donor should prefer $v = 100$ as they can

achieve expected impact of 2. But, if restricted to only 2 samples the performance in the non-perfect scenarios (1 to 4) degrades significantly for higher values of v - so much so that it is optimal for the donor to prefer lower risk levels around $v = 10$. For most of the scenarios the optimal risk level increases significantly as the number of samples increases. However, it remains quite low in all cases for the 'Relatively clueless' scenario.

These results illustrate that it is possible for this sampling constraint to make it optimal for the donor to prefer lower values of v , despite higher values of v offering more upside under ideal conditions. This is a particular form of risk aversion - the next two subsections explore other features that result from the sampling restriction.

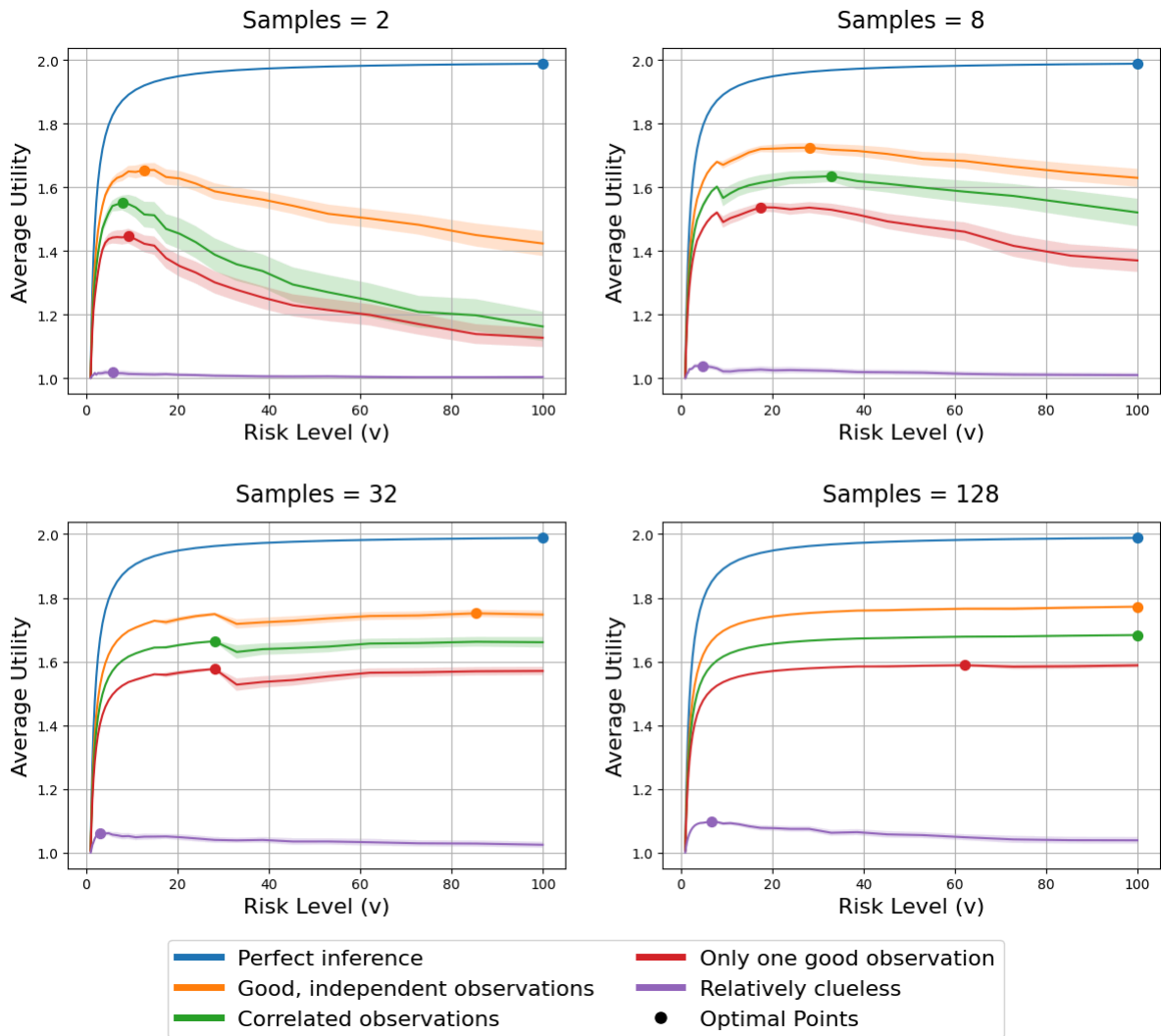


Figure 4: Expected impact-risk level frontier curves for each observation scenario and number of samples. The Expected Impact values are averages across many simulations with two standard deviations of sampling error represented by the shaded areas.

'Ambiguity' Aversion

Uncertainty is often split into two types: risk, where the probabilities of different outcomes are known, and ambiguity (or model uncertainty) where the probabilities themselves are uncertain. The latter isn't explicitly a part of the toy model. However, the sample restrictions can be viewed as producing a similar, arguably equivalent, effect.

Figure 5 reorganizes the results of Figure 4 to make it easy to confirm that as the number of samples increases the expected impact increases, for all values of v . So, it is rational reason for the donor to prefer lower 'ambiguity' (higher sample) situations. Thus, the toy model generates both rational risk aversion (preference for smaller v) and rational ambiguity aversion (preference for situations that are equivalent to having more samples).

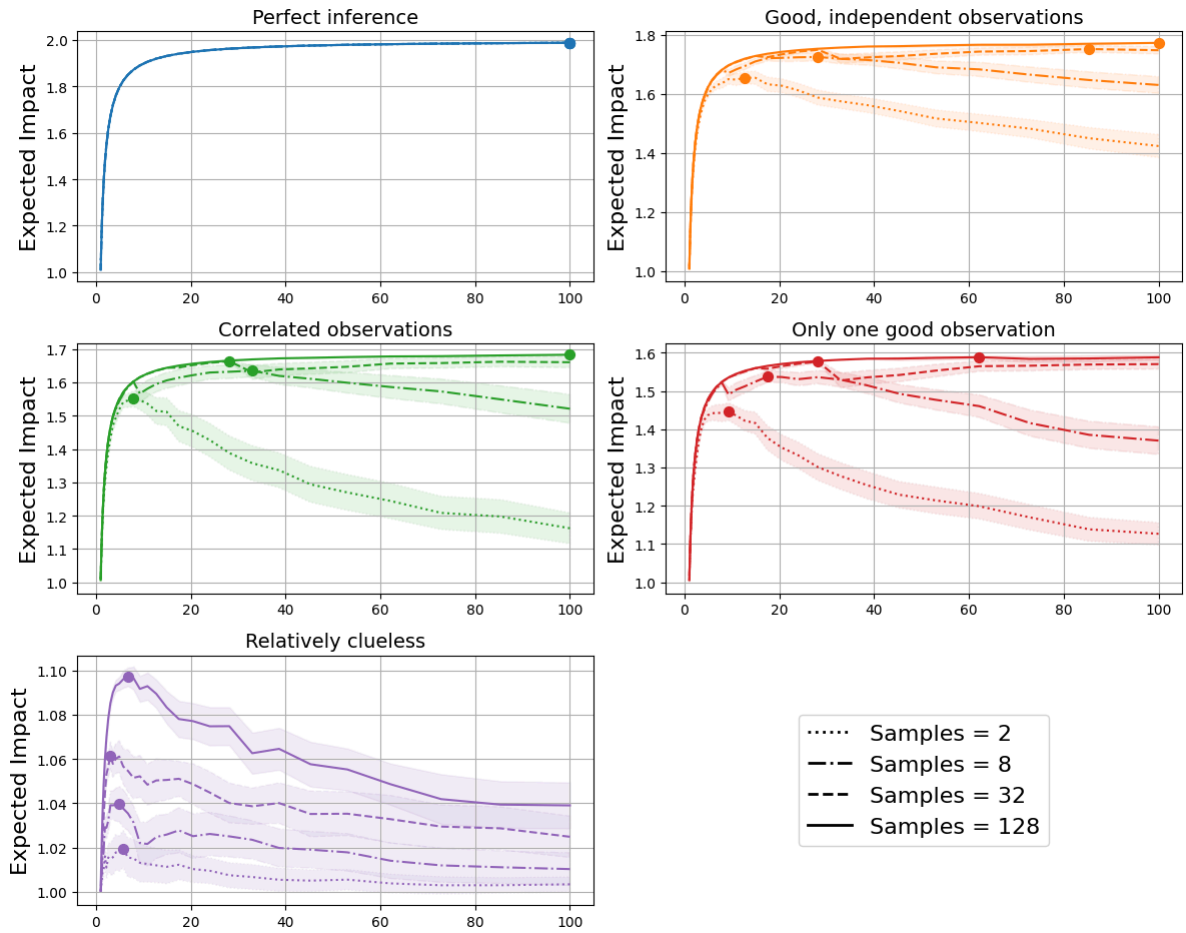


Figure 5: The frontier curves from Figure 4 grouped by observation scenario.

Diversification and Explicit Risk Aversion

While the model above generates uncertainty aversion it doesn't generate diversification. Except for some randomness due to the non-zero bandwidth cost, the agent's samples will either confirm to them that giving everything to A or B is the best policy.

However, the strategy we've allowed the agent is admittedly a bit naive. With a bit more thought they might choose to explicitly, preemptively include risk aversion in their model so they underweight actions with more volatile impacts. This should steer them towards determining their policy based on actions that they have better estimates and mitigate some of the effect of their sampling constraint.

Note that to get such a technique to work, we need to consider situations where the choice between A and B isn't so obvious given the observations. When B is either $s = v = 100$ or $s = 0$, it's almost always going to be clear that is it better or worse than B. We need to consider situations that are more close, so we'll consider $v = 1.1$ and allow the donor 2 samples.

The top chart in Figure 6 shows the frontier curve of expected impact versus impact volatility (the expected standard deviation of the realized impact given repeated trials). It demonstrates that in this case including a bit of preemptive risk aversion is rational in that it improves the expected impact.

The bottom chart in Figure 6 shows for each of the expected impact maximizing points the percentage of time a 'mixed' action of 0.1–0.9 is taken. This shows that rational risk aversion results in diversification even in this binary situation with no inherent risk aversion in the agent's utility function.

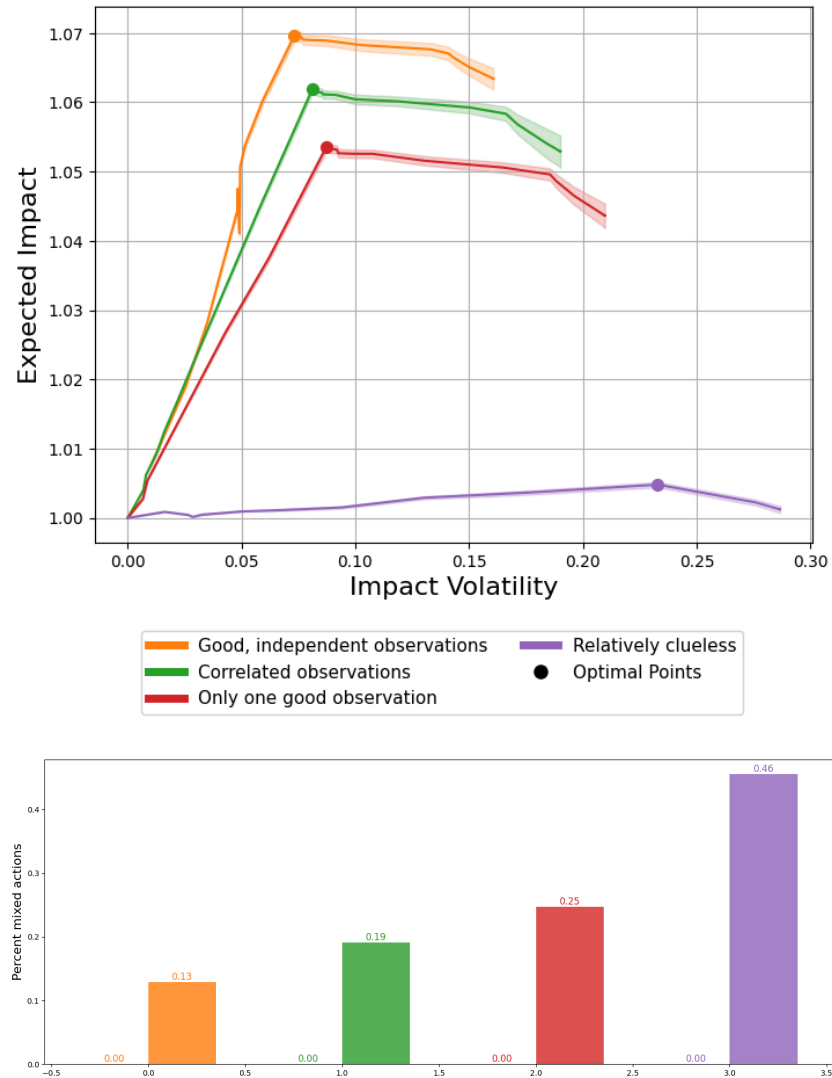


Figure 6: The top chart shows expected impact-volatility frontier curves for each observation scenario with $v=1.1$ and $Samples=2$. The curves are generated by increasing the donor's explicit risk aversion. The Expected Impact values are averages across many simulations with two standard deviations of sampling error represented by the shaded areas. The bottom chart compares the percentage of time a 'mixed' action, $0 < a < 1$, is taken with zero risk aversion (left bars) and with the optimal level of risk aversion (right bars).

Connections To and Implications For Different Fields

This section offers a selection of implications the toy model, and bounded rationality in general, have for different fields.

General

- **Bottom-up paradigm:** Research across related fields is dominated by a top-down paradigm of adding features to models based on heuristic reasoning and seeing what works to improve performance. This toy model, and 'bounded rationality' in general, suggests an alternative 'bottom-up' paradigm of first focusing on better defining the optimization problem including all its constraints. Solving the properly defined problem then naturally leads to optimal performance. In other words, better theory can lead to better models, and the better theory that may be needed is a better description of the practical constraints faced by decision-making agents.
- **Uniting risk and ambiguity aversion:** In the model the different forms of uncertainty aversion all come from the same source of estimation error. This highlights that the distinction between risk and ambiguity aversion may not be as clear as is often assumed in economics and philosophy.

Philanthropy and Impact Investing

- **Open-mindedness:** The inference framing emphasizes the potential benefits of open-minded consideration of different perspectives, especially in challenging, complex decisions like those to do with impact on society.
- **Preference for safer bets:** The toy model doesn't rule out a 'hits-based' approach of making a lot of high-risk, high-reward bets. But the results do point out that in situations of greater uncertainty it can be rational to stick to safer bets.
- **A fundamental reason for diversification:** The model highlights estimation error as a fundamental reason for diversification, whether across charities, investments, strategies, causes, or time. This aligns with the fact that many philanthropic organizations diversify and goes beyond traditional explanations like diminishing-returns to scale, reputation management, and combinatorial effects. The issue of estimation error applies to donors and investors of all sizes and it applies most strongly to more complex and controversial topics like how much to allocate to philanthropy versus impact investing.
- **Frontier curves:** The note highlights the value of visual tools like frontier curves for analyzing optimal policies.

AI Alignment

- **Potential for nuanced AGI behavior and alignment strategies:** While not changing the potential for catastrophic risks from AI, the model suggests advanced AI systems might develop uncertainty aversion and open-mindedness, challenging simplistic views of AI as pure expected utility maximizers. Considering bounded rationality and uncertainty aversion could inform more nuanced AI alignment strategies that don't rely on assumptions of ideal rationality.
- **Potential capability-safety alignment:** If it is optimal for AIs facing complex, highly uncertain situations to consider multiple perspectives and exhibit uncertainty aversion, this could mean that training AIs in a way that acknowledges they are also 'bounded rational' could be a win-win for capabilities and safety in complex situations.
- **Generalized alignment:** If some situations are so complex that both humans and AIs are more or less 'clueless' then a natural question to ask is why focus on AI alignment? It may be more productive to go one level of abstraction up and focus on improving global coordination and robustness in general. This could look like developing coordination mechanisms (including AI training algorithms) that are explicitly aligned with bounded rational inference.

Moral Philosophy

- **Morality as inference:** The model suggests a framework of viewing existing moral theories as imperfect inferences of an ideal theory. The individual theories can then be combined to produce a better approximation of the ideal theory, and this combination need not require the theories to be commensurable. This offers a new perspective on moral uncertainty, potentially informing approaches like 'variance voting' and the 'moral parliament' (MacAskill, Bykvist & Ord, 2020). For example, it suggests that the credences assigned to different moral theories in these existing approaches could be informed by Bayesian calculations based on rough models for how each theory approximates the ideal theory.
- **Fanaticism vs. caution:** The model endogenously generates that in situations of cluelessness safe bets are preferred over high-risk, 'fanatical' bets, while the opposite can be true in situations that are closer to ideal.
- **Inseparability of empirical uncertainty:** The model highlights that adjustments for empirical uncertainty cannot necessarily be treated separately from moral uncertainty, as the optimal policy depends on both analyst models and under bounded rationality this dependence is complex.
- **Challenging the focus on ideals:** This model challenges moral theories that assume unlimited reasoning capabilities and extreme certainty.
- **Computational Moral Philosophy:** It shows that it is possible to integrate insights from neuroscience and machine learning into a philosophically relevant thought experiment that can be analyzed on a computer. Offers mathematical version of moral uncertainty that complements existing more wordy work on 'global consequentialism' and such. Challenges idealized moral theories by integrating insights from neuroscience and machine learning into a computationally analyzable thought experiment.

Techno-philosophies

- **Alignment with d/acc:** The results of the model are perhaps aligned with Vitalik Buterin's d/acc (defensive accelerationism) over the extreme optimism of e/acc and the extreme pessimism of doomers. This is because it highlights that humility and risk aversion can be optimal in the most complex, uncertain situations (like advancing global progress), though it fully supports aggressive, high-risk bets in certain situations.

Economics, Statistics and Machine Learning

- **Endogenous robustness:** This model demonstrates how uncertainty aversion can arise endogenously and rationally, contrasting with approaches that treat robustness as an intrinsic good.
- **Optimal robustness:** By potentially inferring the estimation abilities of different actors, this approach might help define optimal levels of robustness rather than leaving it as a subjective hyperparameter.
- **Adversarial framing:** The model challenges the asymmetric presumption of adversariality in some robustness models, suggesting a more balanced approach to uncertainty.
- **Exploration and uncertainty aversion:** This framework shows how agents can be both uncertainty averse and explorative, reconciling seemingly contradictory aspects of decision-making under uncertainty.

Research Ideas and Questions

This model highlights several promising directions for future research, including:

- Extending the model to address its limitations and to capture more complex, multi-period decision scenarios, and applying it to classic problems in decision theory and economics.
- Investigating how different forms of bounded rationality interact, particularly in multi-agent competitive or cooperative settings.
- Exploring the moral philosophical implications regarding topics like cluelessness, open-mindedness and fanaticism.
- Developing a unified framework that reconciles different types of uncertainty (e.g., risk, ambiguity, Knightian uncertainty) within the context of bounded rationality and computational constraints.
- Developing practical tools and heuristics based on these insights for real-world decision-making in fields such as finance, policy-making, and AI development.

Conclusion

This research note introduces a simple toy model that highlights the important implications of bounded-rational inference for decision-making, including that:

- Combining multiple models can lead to superior outcomes.
- Uncertainty aversion, including both risk and ambiguity aversion, can emerge as a rational strategy in scenarios where perfect information and reasoning is far from possible.
- Frontier curves provide a powerful way to visualize complex decision-making trade-offs.

These insights challenge conventional thinking across multiple disciplines. This includes offering a new perspective on topics like diversification in philanthropy and impact investing, capabilities versus alignment trade-offs in AI, moral uncertainty in philosophy, and robustness in economic models.

I invite researchers from AI, economics, philosophy, and beyond to explore these ideas further. If you find this model intriguing and see potential applications in your field, I'd love to hear from you. Please reach out, whether you're interested in integrating these ideas into your work or developing a paper or applied project together.

As we face increasingly complex global challenges, further study of models like this can help develop a more nuanced understanding of rational decision-making – one that embraces our limitations while striving for optimal outcomes. Future research should focus on refining these models and developing practical applications to enhance decision-making strategies in an uncertain world.

Code

This section presents the code for this example. It was developed by adapting code from Genewein et al. (2015). The full codebase, including this notebook and all supporting files, is available on GitHub [here](#). Readers are encouraged to explore, run, and build upon this code.

```
In [ ]: #Import necessary functions
import Modules.Rational_Frontiers_imports
import importlib
importlib.reload(Modules.Rational_Frontiers_imports)
from Modules.Rational_Frontiers_imports import *

In [ ]: # Generate results for different complexity costs
beta_ao_values = np.logspace(np.log10(.01), np.log10(1000), num=50) # Define the range for
```

```

results_b, scenarios_b = generate_results(da=1,v_values=[2],beta_ao_values=beta_ao_values)

In [ ]: # Generate results for different risk levels and sample sizes
v_values = np.logspace(np.log10(1.01), np.log10(100), num=30) # Define the range for v
results_v, scenarios_v = generate_results(da=0.1,v_values=v_values,n_samp_values=[2,8,32,12])

In [ ]: # Generate results for different levels of explicit risk aversion
g_values = np.concatenate(([0], np.logspace(np.log10(0.0001), np.log10(1), num=30))) #Define
results_g, scenarios_g=generate_results(da=0.1,v_values=[1.1],n_samp_values=[2],gamma_value

In [ ]: # # Save res_v
# with open('Results/results_v.pkl', 'wb') as f:
#     pickle.dump({'results_v': results_v, 'scenarios_v': scenarios_v}, f)

# Load res_v from the pickle file
with open('Results/results_v.pkl', 'rb') as f:
    loaded_data = pickle.load(f)
# Access the loaded variables
results_v = loaded_data['results_v']
scenarios_v = loaded_data['scenarios_v']

In [ ]: # # Save res_g
# with open('Results/results_g.pkl', 'wb') as f:
#     pickle.dump({'results_g': results_g, 'scenarios_g': scenarios_g}, f)

# Load res_g from the pickle file
with open('Results/results_g.pkl', 'rb') as f:
    loaded_data = pickle.load(f)
# Access the loaded variables
results_g = loaded_data['results_g']
scenarios_g = loaded_data['scenarios_g']

In [ ]: # Plot frontier curves for policy complexity vs expected utility
plot_frontiers(results_b,scenarios_b,x_field='I_ao',y_field='E_U',x_label='Policy Complexit

In [ ]: # Plot results for different risk levels and sample sizes
plot_results_v(results_v,scenarios_v)
plot_results_v_byscenario(results_v,scenarios_v)

In [ ]: # Plot frontier curves for impact volatility vs expected utility with explicit risk aversion
plot_frontiers(results_g,scenarios_g,x_field='Vol_U_sample',y_field='E_U_sample',x_label='I
bar_by_scenario(results_g)

In [ ]: # Plot probability models for a specific scenario
plot_scenario_prob_dist(scenarios_v,scenario_id=1,beta_ao=.1)

```

Technical Discussion

Consider a decision making as inference context where an agent uses observations o to infer the world state s and decide a policy $p(a|o)$ for actions a that will maximize their expected utility:

$$\arg \max_{p(a|o)} \mathbf{E}_{p(s,o,a)}[U(s, a)]$$

Information-theoretic frameworks like 'bounded rationality' (Genewein et al., 2015) and 'policy compression' (Lai, Gershman, 2024) add a term to account for the computational cost of converting observations into actions:

$$\arg \max_{p(a|o)} \mathbf{E}_{p(s,o,a)}[U(s, a)] - \frac{1}{\beta_{ao}} I(O; A),$$

where $I(O; A)$ is the mutual information between the distribution of observations and actions. Note that in general they may also add other terms, including $I(O; S)$.

The solution is given by:

$$p^*(a|o) = \frac{1}{Z(o)} p(a) \exp(\beta_{ao} E[U(s, a)|a, o]) \quad (1)$$

$$E[U(s, a)|a, o] = \sum_s p(s|o) U(s, a) \quad (2)$$

$$p(a) = \sum_{s, o} p(s) p(o|s) p^*(a|o), \quad (3)$$

where $Z(o)$ denotes the corresponding normalization constant. In most cases it is not possible to solve these equations analytically. But, generally the solution can be found numerically via the Blahut-Arimoto algorithm.

However, the above solution assumes that the agent can accurately calculate $E[U(s, a)|a, o]$. If this expectation is complex and costly to calculate then the agent will only have a noisy estimate. Even unbiased noise inside the exponential term in the solution can result in suboptimal policies. The simple model in this note assumes the agent makes a noisy estimate of $\hat{E}[U(s, a)|a, o]$ and then repeatedly uses this during the Blahut-Arimoto algorithm.

The model with explicit risk aversion is a modification to:

$$p^*(a|o) = \frac{1}{Z(o)} p(a) \exp\left(\beta_2 \hat{E}[U(s, a)|a, o] - \gamma \text{Var}[U(s, a)|a, o]\right) \quad (4)$$

where γ is the risk aversion parameter and $\text{Var}[U(s, a)|a, o]$ is the variance of U conditional on a and o . This assumes that the agent is endowed with estimates of $\text{Var}[U(s, a)|a, o]$. This may seem unlikely as the key idea here is that the agent struggles to estimate $E[U(s, a)|a, o]$. But, it is plausible the agent has some intuitive idea of the level of $\text{Var}[U(s, a)|a, o]$. Future research could make this model more realistic.

References

- Bhui R., Lai L., & Gershman S. J. (2021) Resource-rational decision making. *Current Opinion in Behavioral Sciences*, Volume 41.
- Buchak, L. How Should Risk and Ambiguity Affect Our Charitable Giving? *Utilitas*. 2023;35(3):175-197.
- Genewein, T., Leibfried, F., Grau-Moya, J. & Braun, D. A. (2015) Bounded Rationality, Abstraction, and Hierarchical Decision-Making: An Information-Theoretic Optimality Principle. *Front. Robot. AI* 2:27.
- Lai, L., Gershman, S. J. (2024) Human decision making balances reward maximization and policy compression. *PLOS Computational Biology* 20(4).
- Lieder, F., Griffiths, T. L. (2019) Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behav Brain Sci*.
- Lieder, F., Hsu, M., & Griffiths, T. L. (2014). The high availability of extreme events serves resource-rational decision-making. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 36.
- MacAskill, W., Bykvist, K., & Ord, T. (2020). *Moral uncertainty*. Oxford University Press.