

Part 1: Mars Rover Routes and Utility

Sandy terrain 3 km/h

Smooth terrain 5 km/h

Rocky terrain 2 km/h

Route 1 is 2 km long. 20% chance sandy, 30% chance smooth, 50% chance rocky

Route 2 is 1.8 km long. 40% chance sandy, 20% chance smooth, 40 % chance rocky

Route 3 is 3.1 km long. 50% chance sandy, 40% chance smooth, 10% chance rocky

(10 points) Which route should we pick? Show your work.

$$\text{Route 1} = .2\left(\frac{2}{3}\right) + .3\left(\frac{2}{5}\right) + .5\left(\frac{2}{2}\right) = 0.753 \text{ hrs}$$

$$\text{Route 2} = .4\left(\frac{1.8}{3}\right) + .2\left(\frac{1.8}{5}\right) + .4\left(\frac{1.8}{2}\right) = 0.672 \text{ hrs}$$

$$\text{Route 3} = .5\left(\frac{3.1}{3}\right) + .4\left(\frac{3.1}{5}\right) + .1\left(\frac{3.1}{2}\right) = 0.920 \text{ hrs}$$

Route 2 is the best.

Route 1 contains a crater. If the wall of the crater is intact, we can go through it. If the wall has been damaged, we will need to go around, which will add 45 minutes to our journey. There is a 30% chance that the wall is damaged.

Route 2 contains a bridge. If that bridge is damaged, we will need to repair it, which will add 1 hour to our time. There is a 60% chance that the bridge is out.

(10 points) Now which route should we pick? Show your work.

$$\text{Route 1} = .2\left(\frac{2}{3}\right) + .3\left(\frac{2}{5}\right) + .5\left(\frac{2}{2}\right) + .3\left(\frac{45}{60}\right) = 0.978 \text{ hrs}$$

$$\text{Route 2} = .4\left(\frac{1.8}{3}\right) + .2\left(\frac{1.8}{5}\right) + .4\left(\frac{1.8}{2}\right) + .6(1) = 1.272 \text{ hrs}$$

$$\text{Route 3} = .5\left(\frac{3.1}{3}\right) + .4\left(\frac{3.1}{5}\right) + .1\left(\frac{3.1}{2}\right) = 0.920 \text{ hrs}$$

Route 3 is now the fastest with the new information.

(10 points) Now suppose that we can use a satellite to find out whether the terrain in route 3 is smooth. Is this helpful? What is the value of this information? Expressed differently, how long are we willing to wait for this information from the satellite?

$$\text{Route 3} = .5\left(\frac{3.1}{3}\right) + .4\left(\frac{3.1}{5}\right) + .1\left(\frac{3.1}{2}\right) = 0.920 \text{ hrs}$$

$$\text{Route 3 | Smooth} = .4\left(\frac{3.1}{5}\right) + (\text{Route 3 w/o Smooth}) = 0.920 \text{ hrs}$$

$$\text{Route 3 w/o Smooth} = .5\left(\frac{3.1}{3}\right) + .1\left(\frac{3.1}{2}\right) = 0.672 \text{ hrs}$$

This information is not helpful. If we know that route 3 is smooth, we simply take route 3 because this is the fastest possible route terrain. If we did not know route 3 was smooth, we would still have taken route 3 because it was already the fastest route. We should not bother waiting for information since it's part of the same option and would not affect our action.

If the best alternative route were not route 3, then we would calculate the value of the information as the new expected value minus the old best expected value.

(5 points) Now put this problem into ChatGPT. Is it able to solve it correctly? If not, where does it make mistakes?

ChatGPT is not able to solve the problem correctly. It does the utility calculation setup correctly, but it does not compute the values correctly. For example, " $0.2 * (2 \text{ km} / 3 \text{ km/h}) = 0.4 \text{ hours}$ " in its Route 1 calculation is incorrect, as it should be 0.13. In the followup with the obstacles in Route 1 and Route 2, the calculations were done correctly but overall was still incorrect. ChatGPT did not state a conclusion for part 1, and for part 2 it has wording which implies Route 2 was not the best choice in part 1 (which it is with its incorrect calculations). The explanation for part 3 is technically correct but does not actually solve it.

ChatGPT's trimmed answer below:

For Route 1:

- Expected time on sandy terrain = $0.2 * (2 \text{ km} / 3 \text{ km/h}) = 0.4 \text{ hours}$
- Expected time on smooth terrain = $0.3 * (2 \text{ km} / 5 \text{ km/h}) = 0.6 \text{ hours}$
- Expected time on rocky terrain = $0.5 * (2 \text{ km} / 2 \text{ km/h}) = 1 \text{ hour}$
- Total expected time for Route 1 = $0.4 + 0.6 + 1 = 2 \text{ hours}$

For Route 2:

- Expected time on sandy terrain = $0.4 * (1.8 \text{ km} / 3 \text{ km/h}) = 0.24 \text{ hours}$
- Expected time on smooth terrain = $0.2 * (1.8 \text{ km} / 5 \text{ km/h}) = 0.36 \text{ hours}$
- Expected time on rocky terrain = $0.4 * (1.8 \text{ km} / 2 \text{ km/h}) = 0.72 \text{ hours}$
- Total expected time for Route 2 = $0.24 + 0.36 + 0.72 = 1.32 \text{ hours}$

For Route 3:

- Expected time on sandy terrain = $0.5 * (3.1 \text{ km} / 3 \text{ km/h}) = 0.5167 \text{ hours}$
- Expected time on smooth terrain = $0.4 * (3.1 \text{ km} / 5 \text{ km/h}) = 0.744 \text{ hours}$
- Expected time on rocky terrain = $0.1 * (3.1 \text{ km} / 2 \text{ km/h}) = 0.155 \text{ hours}$
- Total expected time for Route 3 = $0.5167 + 0.744 + 0.155 = 1.4157 \text{ hours}$

Now, let's consider the additional information:

For Route 1:

- 30% chance the wall of the crater is damaged, adding 45 minutes (0.75 hours) to the journey.
- Total expected time for Route 1 with obstacle = 2 hours + (0.3 * 0.75 hours) = 2.225 hours

For Route 2:

- 60% chance the bridge is damaged, adding 1 hour to the journey.
- Total expected time for Route 2 with obstacle = 1.32 hours + (0.6 * 1 hour) = 1.92 hours

With this new information, Route 2 becomes the better choice because the expected time for Route 2 with the obstacle is still less than Route 1 with the obstacle.

Regarding Route 3 and using a satellite to find out if the terrain is smooth, the value of this information depends on how long you are willing to wait for it. You should compare the expected time for Route 3 (1.4157 hours) with the expected time for other routes after factoring in the potential obstacles. If waiting for the satellite information and then choosing Route 3 still results in a faster overall journey, then it would be helpful. Otherwise, it may not be worth waiting for the satellite data.

Part 4: AINow

What are the three dimensions along which Big Tech has an advantage in AI?

- Data - Companies have endless amounts of behavioral data collected about customers (especially with the consolidation of companies through acquisition) to make AI products.
- Computing Power - AI needs a lot of computing power to train/tune/deploy which is expensive, and big tech owns a lot of cloud/compute infrastructure that are used by other companies (which means they can throttle/restrict them).
- Geopolitical - Rhetoric of the US-China AI race has positioned big tech as equivalent to a US power, allowing companies to grow further and gain stronger political standing.

Why does AI Now think it's important to focus on Big Tech?

- Addressing issues which are exemplified by big tech companies can address the root causes (e.g. invasive data surveillance, consolidation of economic power, exacerbation of inequity)
- Business and regulatory actions taken by big tech companies influence other companies to do so as well.
- With greater consolidation of power, companies become a chokepoint and single point of failure (e.g. government dependence on big tech infrastructure)

Priority 1 discusses Algorithmic Accountability. What does this mean? Why is it important to shift responsibility for detecting harm on companies themselves?

- Algorithmic Accountability is about positioning the burden on tech companies to identify and address tech-enabled harms. A due-diligence approach requiring firms to comply with the law, rather than relying on 3rd party regulators, is similar to how other goods are regulated. This also aids in proactively reducing harm in a “move fast and break things” tech environment.

What are the windows for action that are identified? Which do you personally think are the most effective or promising?

- Of the windows for action, data minimization, strong antitrust laws to prevent consolidation of data and other abuses, and maintaining algorithmic transparency are what I find most effective.
- Biometrics and Algorithmic Management are issues but less significant when trying to reduce the breadth and power of big tech. For Algorithmic Accountability, I disagree with displacing the burden of audits to companies. In the current tech environment, companies aren't being held accountable in the fast-moving world of tech. Companies are already providing their own rules/ethics but do not adhere to them themselves. Without strong government regulation and third-party audits, companies have no incentive to self-govern.

The executive summary contains this quote:

"These are only a handful of examples, and what they make clear is that there is nothing about artificial intelligence that is inevitable. Only once we stop seeing AI as synonymous with progress can we establish popular control over the trajectory of these technologies and meaningfully confront their serious social, economic, and political impacts—from exacerbating patterns of inequality in housing, credit, healthcare, and education to inhibiting workers' ability to organize and incentivizing content production that is deleterious to young people's mental and physical health."

Do you agree with this assessment? How might we rethink our relationship with AI and with technology in order to avoid these potential negative outcomes?

I agree with the harmful impacts of AI but am not clear on what they mean by "establish popular control over the trajectory of these technologies" and the extent which they refer to. In general, we could "control" the trajectory of AI by creating stricter regulations and government oversight over big tech to proactively reduce the amount of harm possible. As is, companies are economically incentivized to move fast, develop new products, and maximize engagement over the human aspect; companies creating AI need to be held to a human-first, ethical-AI standard with strong algorithmic transparency measures to minimize the possible harm and maximize utility.

Now paste this last question into ChatGPT and include its response. How do you compare its output to your own?

ChatGPT interprets/summarizes the quote accurately to how I did and provides some response to each of the mentioned areas of impact, whereas I focused on only a few more broad ones to avoid redundancy.

ChatGPT's answer below:

As an AI model, I don't have personal opinions, but I can provide insights into the statement and offer some considerations.

The statement emphasizes the importance of not viewing artificial intelligence as an inevitable force for progress and highlights the potential negative social, economic, and political impacts of AI technologies. It's crucial to recognize that AI technologies are not inherently good or bad; their impact depends on how they are developed, deployed, and regulated.

To avoid potential negative outcomes associated with AI, society can rethink its relationship with AI and technology in several ways:

1. **Ethical AI Development:** Prioritize ethical considerations in the development of AI systems. This includes addressing biases in AI algorithms, ensuring transparency and fairness, and involving diverse stakeholders in the decision-making process.
2. **Regulation and Oversight:** Governments and regulatory bodies should establish clear guidelines and regulations for the development and use of AI technologies. This can help prevent the misuse of AI in areas like housing, credit, healthcare, and education.
3. **Education and Awareness:** Promote public awareness and education about AI and its potential impacts. Informed citizens and consumers can make better choices and advocate for responsible AI practices.
4. **Social Responsibility:** Encourage tech companies to adopt social responsibility in their AI development. This includes considering the broader societal implications of their products and services.
5. **Collaborative Approach:** Engage multiple stakeholders, including government, industry, academia, and civil society, in discussions and decision-making related to AI. Collaboration can lead to more comprehensive and balanced solutions.
6. **Economic Redistribution:** Address the potential economic impacts of AI, such as job displacement, by exploring mechanisms for economic redistribution, retraining, and upskilling of workers.
7. **Mental Health and Well-being:** Promote the responsible use of AI in content production, especially concerning content that affects mental and physical health. Encourage platforms to monitor and address harmful content.
8. **Privacy Protection:** Strengthen data privacy laws and empower individuals to have more control over their personal data, which can help mitigate privacy-related concerns.
9. **Community Engagement:** Encourage local communities to have a say in the deployment of AI technologies in their areas, especially in matters like urban planning and housing.
10. **Long-term Thinking:** Consider the long-term implications of AI development and deployment. What might seem beneficial in the short term can have negative consequences in the future, so it's important to think beyond immediate gains.

In summary, rethinking our relationship with AI and technology involves a multi-faceted approach that combines ethics, regulation, education, collaboration, and societal responsibility. It's essential to recognize the potential risks and actively work to shape the development and use of AI in ways that benefit society as a whole while minimizing negative consequences.