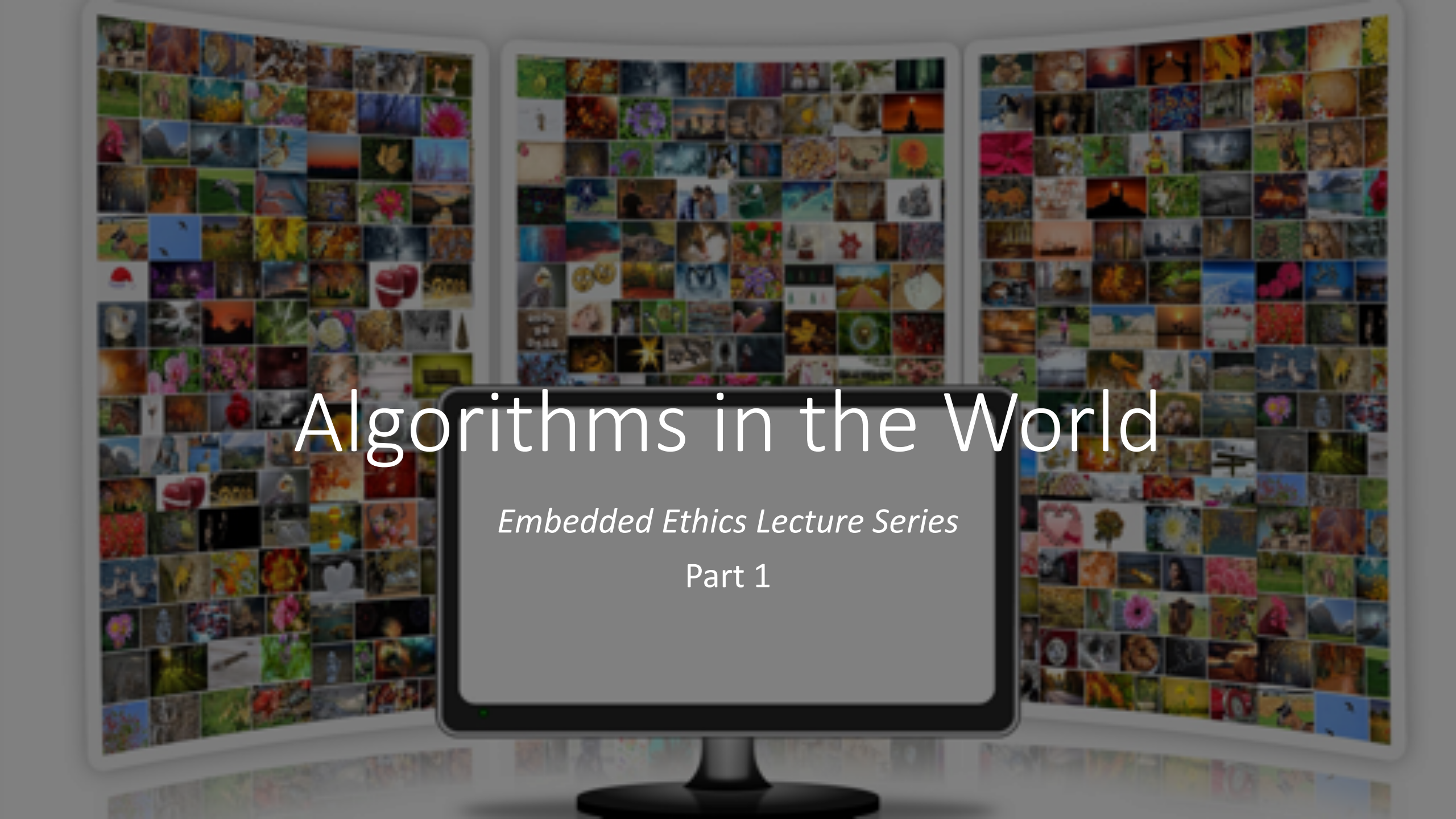
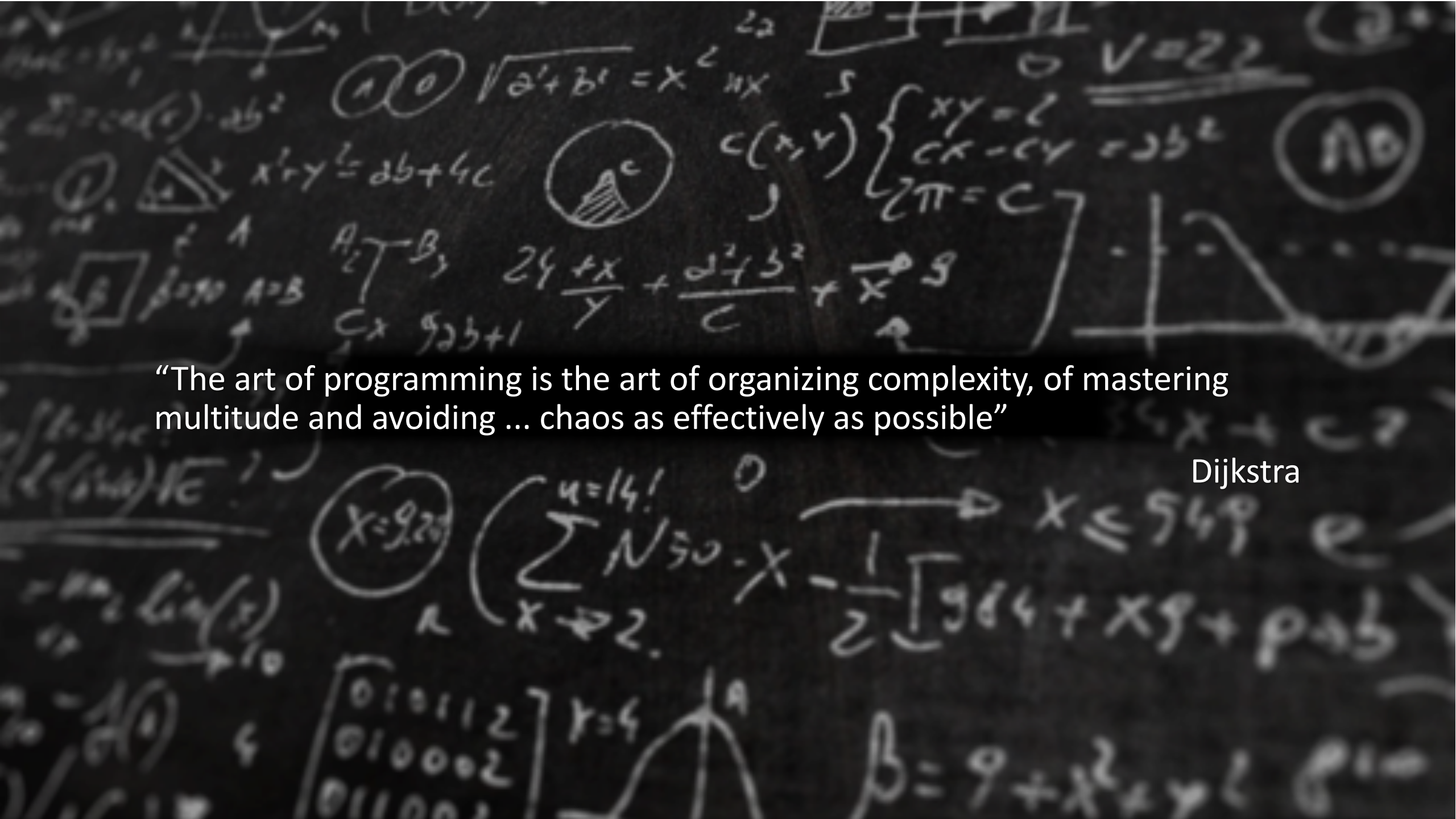


Algorithms in the World

Embedded Ethics Lecture Series

Part 1



“The art of programming is the art of organizing complexity, of mastering multitude and avoiding ... chaos as effectively as possible”

Dijkstra

The world presents a
“blooming, buzzing confusion”

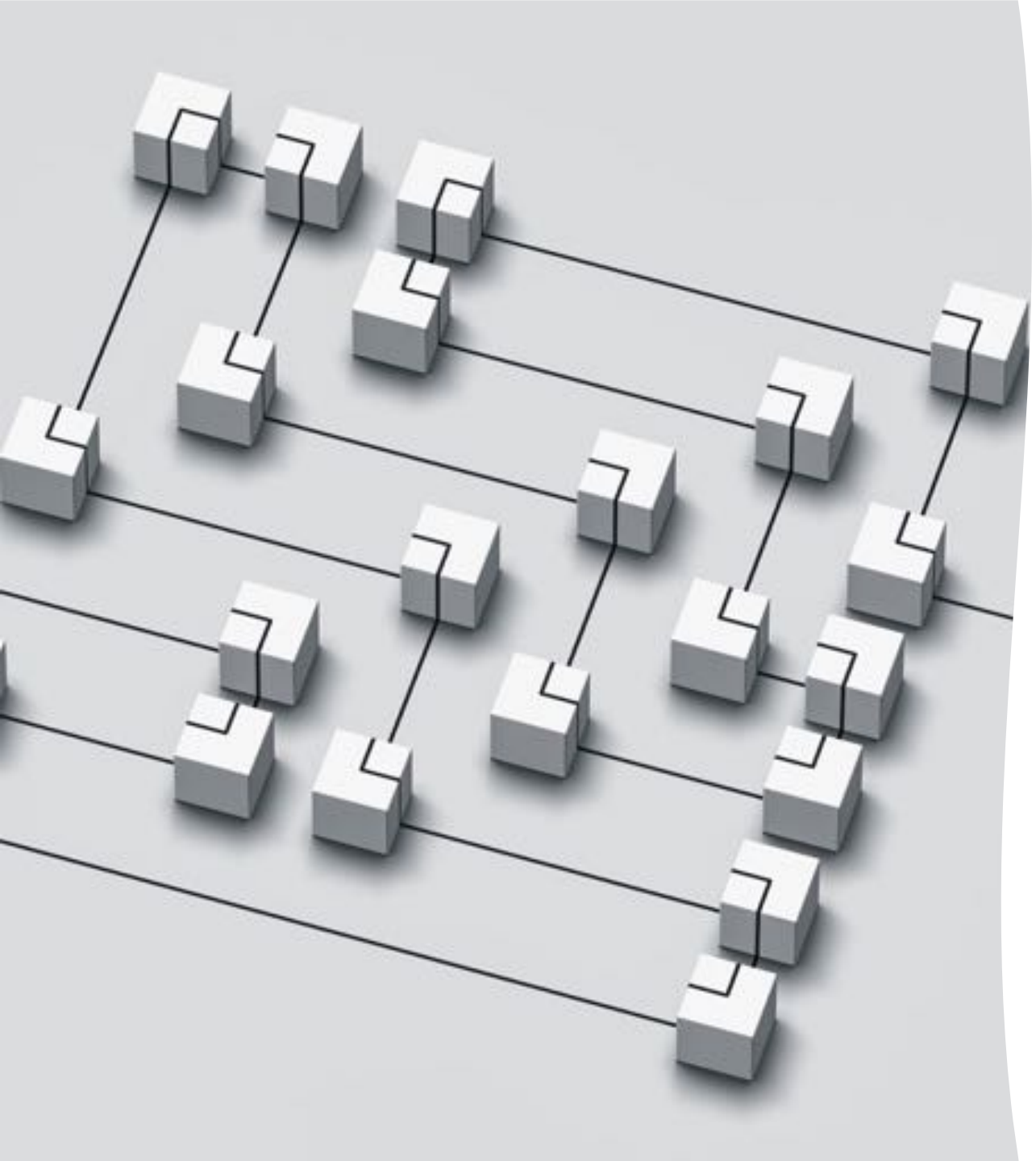
…

… from which we create
algorithmically solvable problems

… and yet the real world sometimes slips out of the grasp of our formalism.

Aspects algorithmic design where the world is not in line with formalism

Decisions with societal impact and ethical dimensions



Contents

Faithful Translation

- Measurement
- Comparison
- Problem Formulation

Faithful Translation

Is this problem a faithful translation of the one we are trying to solve?

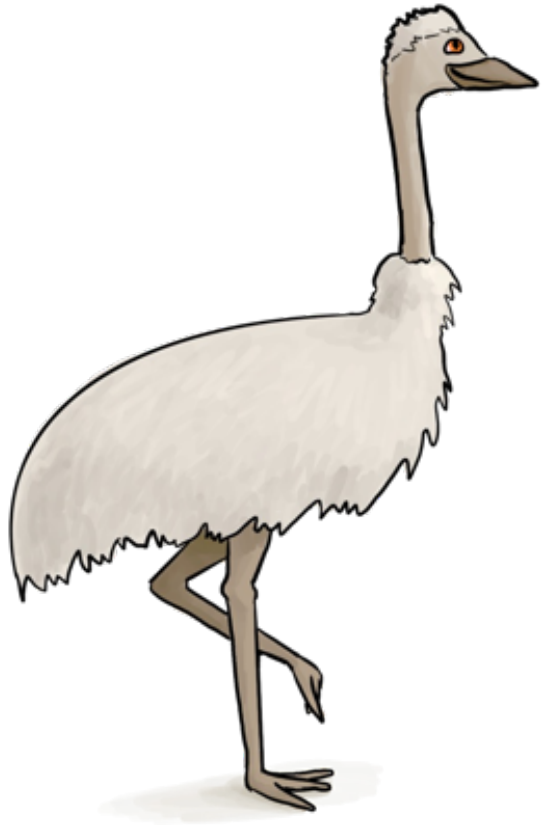
Faithful Translation

[Business problems] must be made into questions that data science can answer. Practitioners are frequently charged with turning amorphous goals into well-specified problems—that is, problems faithful to the original business objectives, but also problems that can be addressed by predicting the value of a variable.

Passi & Barocas, “Problem Formulation and Fairness”

Faithful Translation

- Problem formulation aspires to “faithful translation” between the real-world problem and an algorithmically solvable formalism.
- That is, a correct mapping between corporate goals, organizational data, and computational problems.
- However, the translation is often “negotiated” contingent on the discretionary judgement of various stakeholders and the agreements negotiated between them, and further impacted by the choice of methods, instruments, and data.



ELAN the ETHICAL EMU!

- What kinds of decisions are negotiated?
- Whose perspectives and interests inform these decisions?
- What is the impact of these decisions for key stakeholders?

Measurement

Is what we can measure what we (really) want to measure?

Measurement

- Turning a real-world problem into an algorithmically tractable problem requires a process of formalization.
- One of the first steps in this process consists in identifying features of the real world that are relevant to our problem and “translating” them to quantifiable variables.
- What we **can** measure may act as a proxy for what we **want to** measure
- … but some information may be lost in translation.

Abstraction is when we omit details of the real-world situation.

Omit the kind of thing being sorted by our algorithm, or what condition it is in, or what color it is, or how long it has been in the list.

Idealization is when we deliberately change aspects of the real-world situation.

Round the numbers being sorted to make them whole numbers.

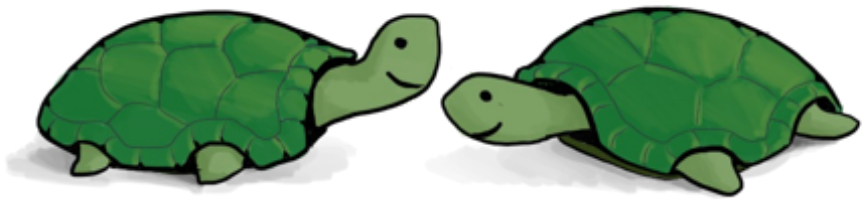
Speed Measurement Case

- Bogotá's Secretariat of Mobility needed to measure average speeds on a road in order to improve their congestion reduction strategy.
- They set up two Bluetooth sensors spaced apart on a roadway. By picking up cell phone identifiers, Bluetooth sensors can determine the speed of the vehicle containing that cell phone.

- What information do we get?
- What information is missing?

- What information do we get?
- What information is missing?

Think-Pair-Share!



Speed Measurement Case

- Bluetooth receptors can only measure speed of vehicles that have Bluetooth or contain people with smart phones.
- They can’t (normally) distinguish between cars, motorcycles, pedestrians, bicycles, and public transit.
- This may have significant effect, as policy may be misguided if it overlooks the information that the algorithm leaves out and treats the outcome as a perfect proxy.

Risks

- Collective risks from misguided policy decisions
- Individual risks
  - Underinclusion in data
  - Overinclusion in data

Risks

- Overinclusion:
  - Violation of privacy
  - Risks to civil liberties
  - Data scraped without valid consent
    - Not informed
    - Coercive conditions

Problematic Overinclusion

- Disproportionate surveillance of welfare and public interest housing recipients
- Conditioning access to basic goods on data sharing.
- Curtailing civil liberties.

Risks

- Underinclusion:
  - Data does not describe you
  - Your needs not measured/accounted for

Problematic Underinclusion

- Lower performance of pulse oximeters for patients of color.
- Higher risk of undetected hypoxemia.

Measurement

- Quantifiable variables may or may not be good proxies for the features of the world that we want to measure.
- Selecting imperfect proxies leads to some relevant information being excluded and some irrelevant information being included.
- Algorithms that employ imperfect proxies may lead to misguided decisions

Comparison

Are two values (truly) comparable?

Comparison

- Recall your sorting algorithm?
- The algorithm's job was to output a correctly sorted list of objects.
- It relied on having answers to questions of the form:

  is [this] bigger than [that]?

- This works very well when the items you are sorting have meaningful values such as…

9	6	3	5	2	1	2

Comparison

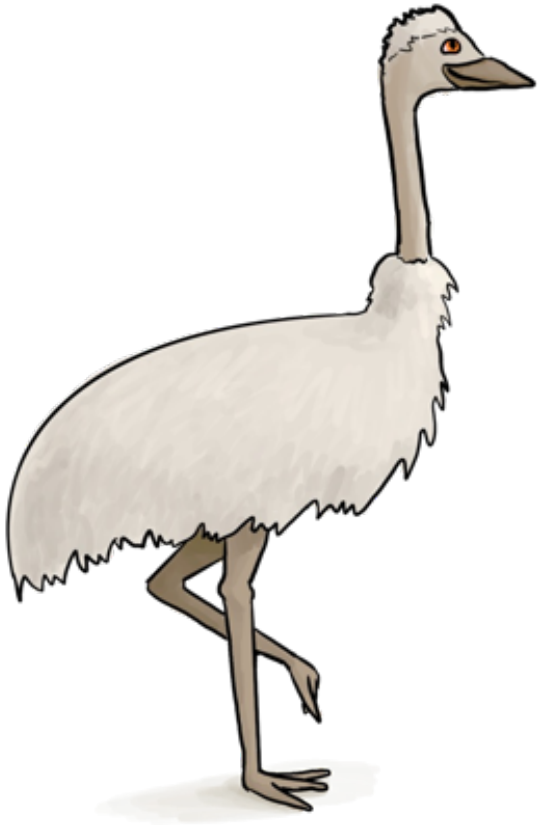
- Recall your sorting algorithm?
- The algorithm's job was to output a correctly sorted list of objects.
- It relied on having answers to questions of the form:

  is [this] bigger than [that]?

- This works very well when the items you are sorting have meaningful values.
- But what happens here?

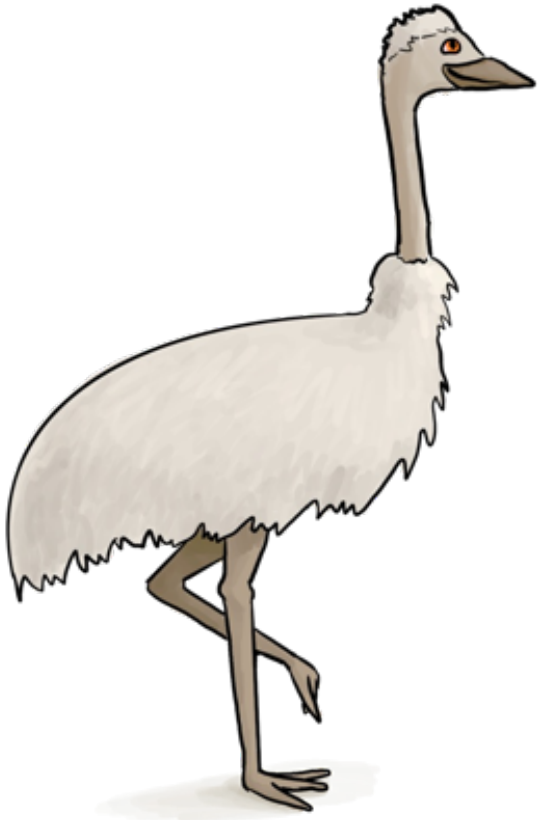
Reliable Travel

- Imagine designing an algorithm for yourself that maximized the reliability of the flights you booked for your summer trips.
- If you consider using this algorithm for an online travel agency, other considerations may become salient.



ELAN the ETHICAL EMU!

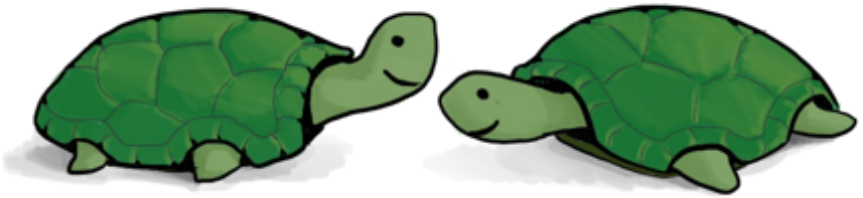
- What other things may be important?
- What are we putting aside?

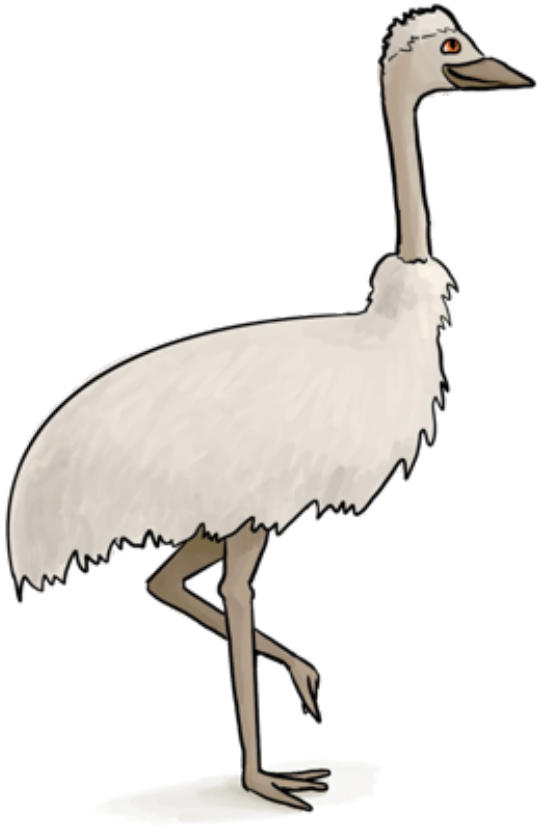


ELAN the ETHICAL EMU!

- What other things may be important?
- What are we putting aside?

Think-Pair-Share!





ELAN the ETHICAL EMU!

- How do we prioritize among these different considerations?

Comparison

Incommensurability: two kinds of things are incommensurable when we lack a *common measure* of value.

Incommensurability makes it difficult to establish ranking relationships.

Reliable Travel

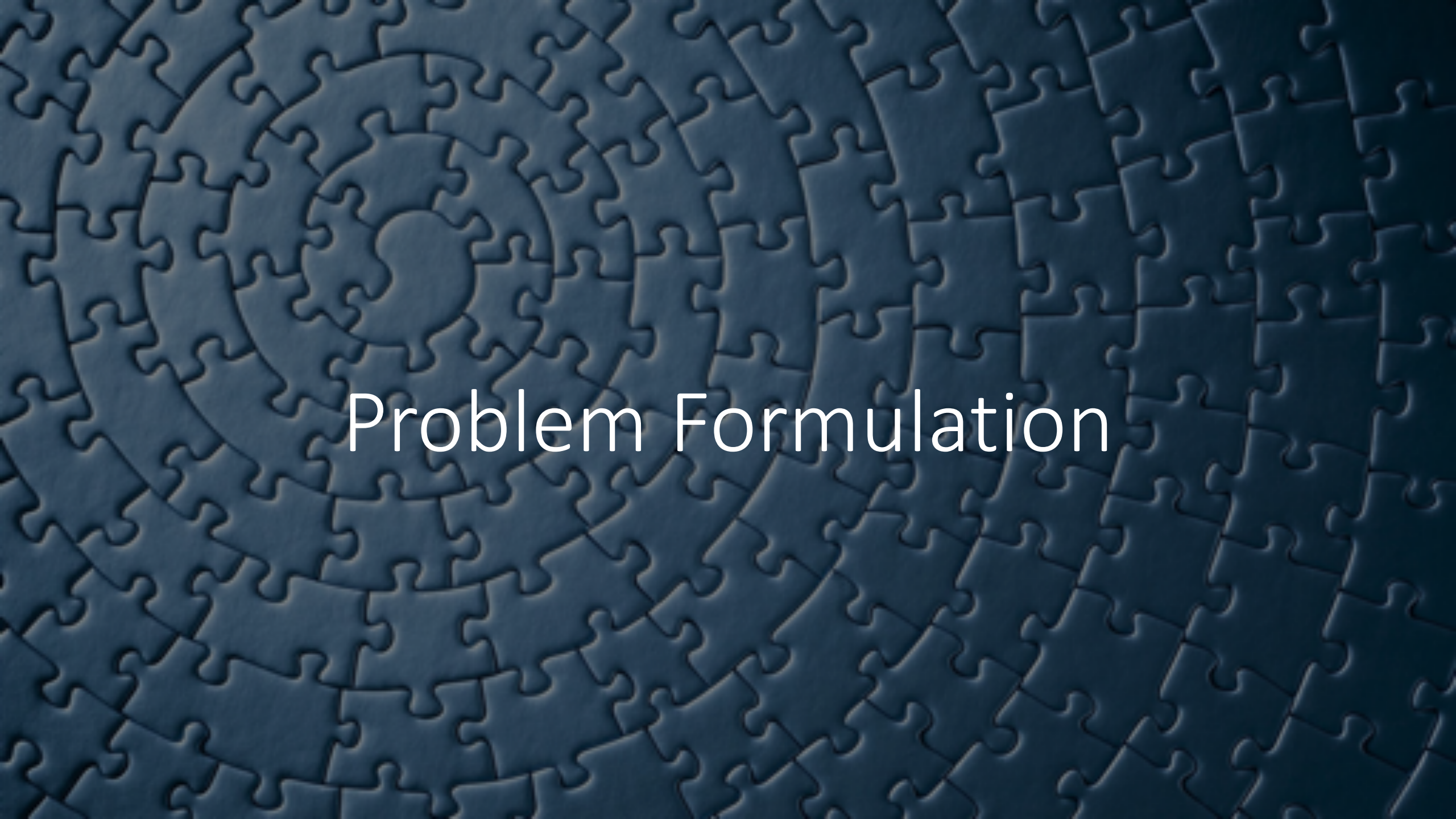
- What do we do when we must account for different values?
- We can rank and sort possible routes on each scale separately. But what if we need to consider both?
- Convert each value into a common measure (like money)
- In that step there is **abstraction** and **idealization.**

Reliable Travel

- As an advisor to the travel agency, you need to make a series of decisions over
  - How you compare different values
  - How you prioritize between them
- Each decision may have significant impact.
- Be sure to make those decisions deliberately and thoughtfully.

Comparison

- Sorting and ranking occur in many contexts.
- To perform these tasks, we need items to have comparable values.
- This is often not the case, as values may be incommensurable.
- Finding a common measure may require abstraction and idealization, which may impact the quality of decisions.
- This may, in turn, impact our collective interests and our interests as individuals.


Problem Formulation

Problem Formulation

Negotiated Translations:

Contingent on the discretionary judgement of various stakeholders and the agreements negotiated between them

Special Financing for Car Purchases: Case Study

- A purchaser wants to buy a good car that they can afford.
- A car dealership wants to sell cars to people who can pay for them and who ideally won’t miss payments or default on a loan.
- A special financing company acts as the intermediary to recommend good “leads” to the dealership.

Special Financing for Car Purchases: Case Study

Business problem: Improving lead quality

- How likely is a given person to be able to finance a car?

Special Financing for Car Purchases: Case Study

What makes a good lead?

- **Dealer Decisions:** How likely is someone to be able to finance a car *at that dealership*?
- **Credit Score:** Maybe a quality lead is one with a credit score over 500?

Special Financing for Car Purchases: Case Study

What makes a good lead?

- **Dealer Decisions:** How likely is someone to be able to finance a car *at that dealership*?
- **Credit Score:** Maybe a quality lead is one with a credit score over 500?

Different problems!

Special Financing for Car Purchases: Case Study

“The quality of a lead is not a preexisting variable waiting to be measured, but an artifact of how our actors define and measure it.”

Passi and Baroccas

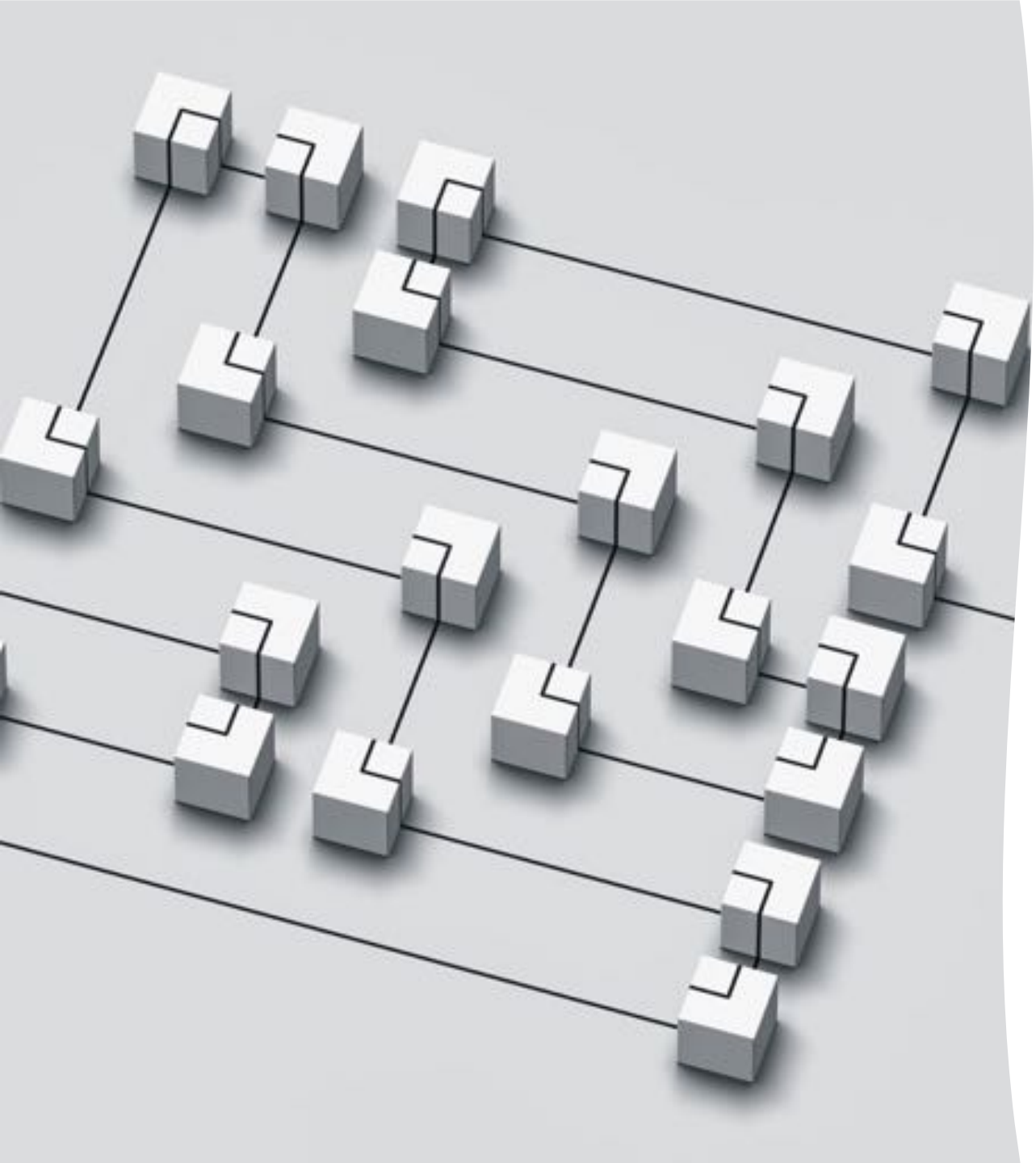
Interests and problem formulation

The clients of special financing companies are the car dealerships. They are solving a problem from the car dealers' perspective.

- How would the problem be formalized if other actors had been involved in the negotiation?
- Who should be involved in the “negotiation” of negotiated translation and algorithmic formalization?

Faithful Translation

- Translating business problems into algorithmically tractable problems requires a process of negotiation, where discretionary judgment plays a substantive role.
- Different formalizations will promote different goals and their societal impact will vary.
- The framing of a problem is often informed by the interests of particular actors.
- It’s important to ask ourselves what the problem might look like if it was framed by other parties’ interests, and who should be at the negotiation table when these decisions are made.



Conclusion

- Three decision points:
  - What do we measure and how?
  - How do we compare things on different value scales?
  - How do we formulate a problem?

- Complexities:
  - Abstraction and idealization
  - Incommensurability
  - Perspectives and interests in negotiating problem translations


Thank you!
dacostan@stanford.edu