Lecture 1: Introduction to Predictive Modeling

Professor Ilias Bilionis

What is predictive modeling?

What can go wrong in an engineering system?

- Engineering systems can be very complicated.
- The are **known knowns.**
- There are **known unknowns.**
- There are **unknown unknowns.**

Orbiter challenger as it lifts from Pad 39A

PREDICTIVE
SCIENCE LABORATORY

A shot at defining predictive modeling

Predictive modeling is the process of describing our state of knowledge about *known unknowns* in order to make informed decisions.

What about unknown unknowns?

- There is no widely accepted automated way for turning unknown unknowns into known unknowns.
- Effectively, this is what science is doing.
- Automating the process requires understanding how humans perform induction.
- It may require general artificial intelligence.

Lecture 1: Introduction to Predictive Modeling

Professor Ilias Bilionis

Causal models and their graphical representation

PREDICTIVE SCIENCE LABORATORY

Why worry about causality?

- Models that explicitly encode causal relationships are the only useful models.
- Most physical and engineering models are causal models.
- When we extend them to account for known unknowns, we need to make sure that they remain causal.
- Otherwise, we are merely capturing correlations and the results cannot be trusted.
- Structural causal models give us the language we need to formalize causality.

What is a structural causal model?

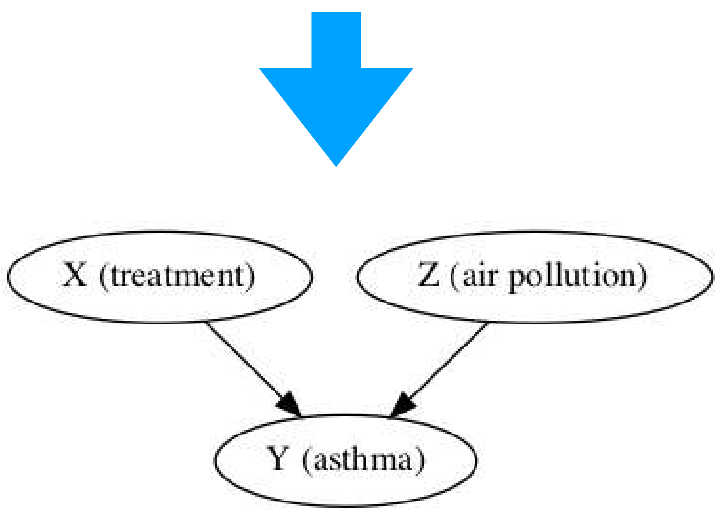
- A set of variables we are interested in:
  - Y: an individual has asthma or not
  - X: the individual is treated or not
  - Z: air pollution level
- A set of functions that describes how the variables are connected:

$$Y = f(X, Z)\,.$$

PREDICTIVE
SCIENCE LABORATORY

Graphical representation of causal models

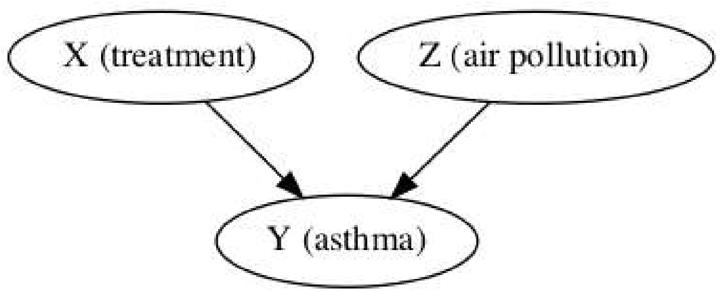
$$Y = f(X, Z)$$



- To each structural causal model there corresponds a directed acyclic graph.
- The edge represents a direct causal link between the parent and the child nodes.

We typically first build the graph and then work out the function details

$$Y = f(X, Z)$$



PREDICTIVE
SCIENCE LABORATORY

Lecture 1: Introduction to Predictive Modeling

Professor Ilias Bilionis

Aleatory vs epistemic uncertainty

PREDICTIVE
SCIENCE LABORATORY

Types of uncertainty

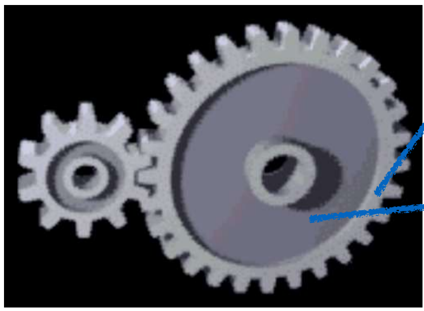
- **Aleatory**: naturally occurring randomness that we cannot (or do not know how to) reduce.

  Latin aleatorius of a gambler, from aleator gambler, from alea a dice game

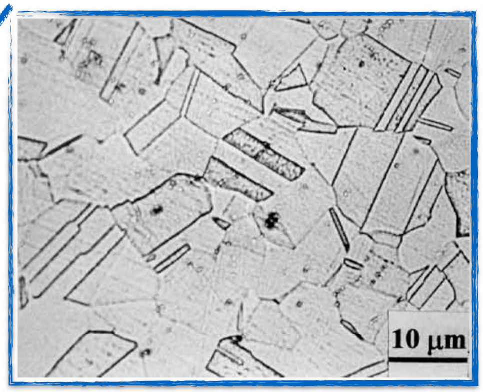
- **Epistemic**: uncertainty due to lack of knowledge that we can reduce by paying a price.

  Greek επιστήμη meaning knowledge.

Unknown microstructure of a manufactured artifact



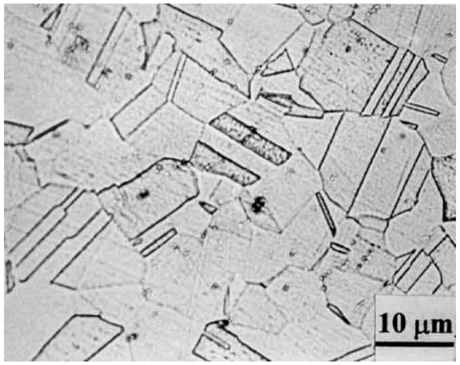
https://www.osha.gov/SLTC/etools/machineguarding/animations/gears.html



https://commons.wikimedia.org/wiki/File:Microstructure_of_a_unsensitised_type_304_stainless_steel.jpg
(https://creativecommons.org/licenses/by-sa/3.0)

We model uncertainties using probability

$p(A \mid K)$ = "How much do we believe A is true given our current state of knowledge K"

p( |K)=?

https://commons.wikimedia.org/wiki/File:Microstructure_of_a_unsensitised_type_304_stainless_steel.jpg
(https://creativecommons.org/licenses/by-sa/3.0)

PREDICTIVE
SCIENCE LABORATORY

Lecture 1: Introduction to Predictive Modeling

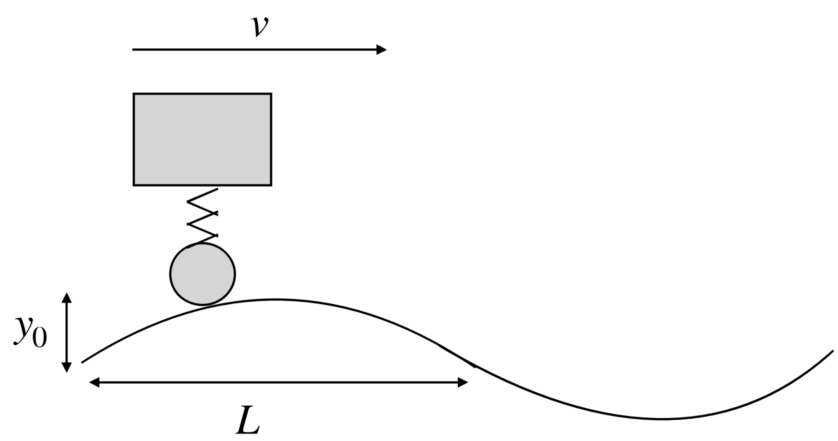
Professor Ilias Bilionis

The uncertainty propagation problem

PREDICTIVE
SCIENCE LABORATORY

Example: Driving a trailer on a bumpy road

- m: mass
- k: spring constant
- v: velocity
- y_0: amplitude of road roughness
- L: "wavelength" of road roughness



<u>Dynamics</u>

$$\omega = \frac{2\pi v}{L}$$

Angular velocity

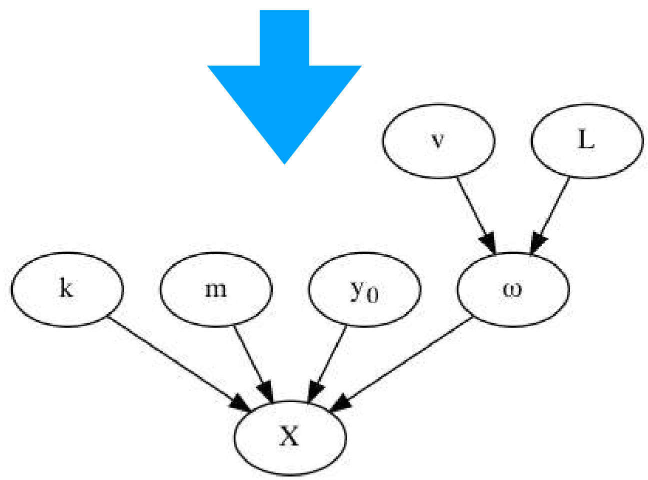
$$X = \left| \frac{k y_0}{k - m\omega^2} \right|$$

Amplitude

PREDICTIVE SCIENCE LABORATORY

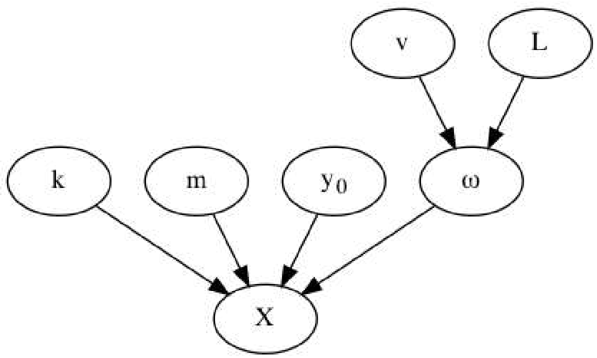
Example: Driving a trailer on a bumpy road

$$\omega = \frac{2\pi v}{L} \qquad X = \left|\frac{k y_0}{k - m\omega^2}\right|$$



PREDICTIVE
SCIENCE LABORATORY

Example: Driving a trailer on a bumpy road



Variable	Type	Values
k	Manufacturing uncertainty	[159,999, 160,001] N/m
v	Operating condition	[80, 150] km/hour
m	Loading condition	[100, 200] kg
y	Road condition	[0, 100] mm
L	Road condition	[1, 2] m

Our state of knowledge about the problem.

PREDICTIVE
SCIENCE LABORATORY

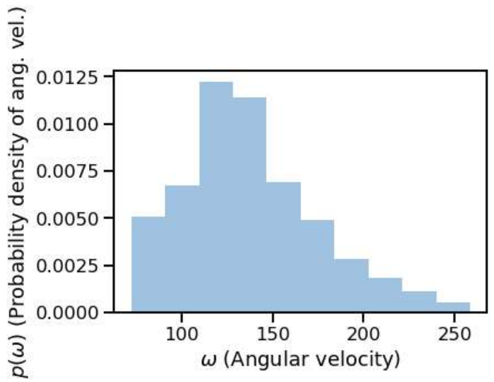
The uncertainty propagation problem

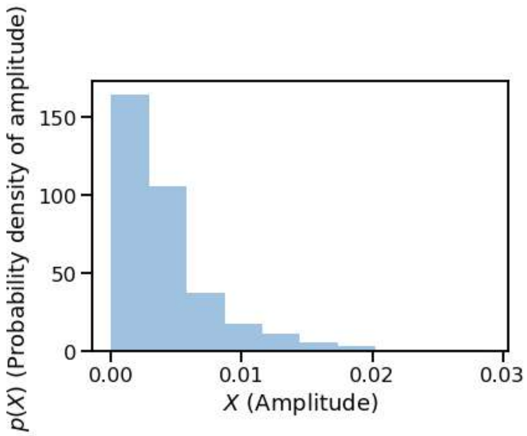
Having quantified our uncertainty about all unknowns, propagate this uncertainty through the causal model to characterize our uncertainty about a quantity of interest.

PREDICTIVE
SCIENCE LABORATORY

The Monte Carlo solution to the uncertainty propagation problem

- Sample random inputs many times.
- Evaluate model outputs at these inputs.
- Estimate any statistics of interest.





PREDICTIVE
SCIENCE LABORATORY

Lecture 1: Introduction to Predictive Modeling

Professor Ilias Bilionis

The model calibration problem

PREDICTIVE
SCIENCE LABORATORY

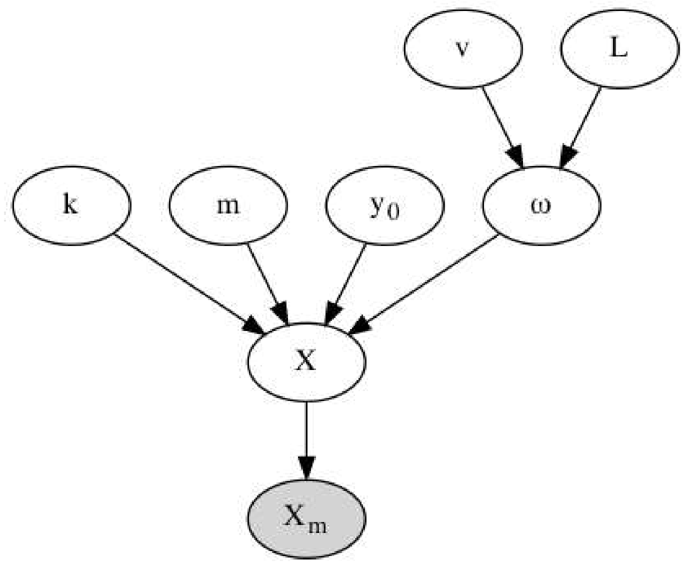
The model calibration problem

- The model calibration problem is the inverse of the uncertainty propagation problem.
- That is why such problems are also called **inverse problems**.
- We observe a quantity that is predicted by the model and we want to characterize how this observation changes our state of knowledge about the model parameters.

PREDICTIVE
SCIENCE LABORATORY

Example: Driving a trailer on a bumpy road

- m: mass
- k: spring constant
- v: velocity
- y_0: amplitude of road roughness
- L: “wavelength” of road roughness
- X_m: the measurement



PREDICTIVE
SCIENCE LABORATORY

The formal solution to the model calibration problem

- Quantify our **prior** state of knowledge about all the model parameters.
- Use Bayes' rule to condition the prior knowledge on the observations to get the **posterior** state of knowledge.
- Create a practical procedure that characterizes our posterior state of knowledge.