BSCS 2019 - Neural Computation

II - Knowledge representations

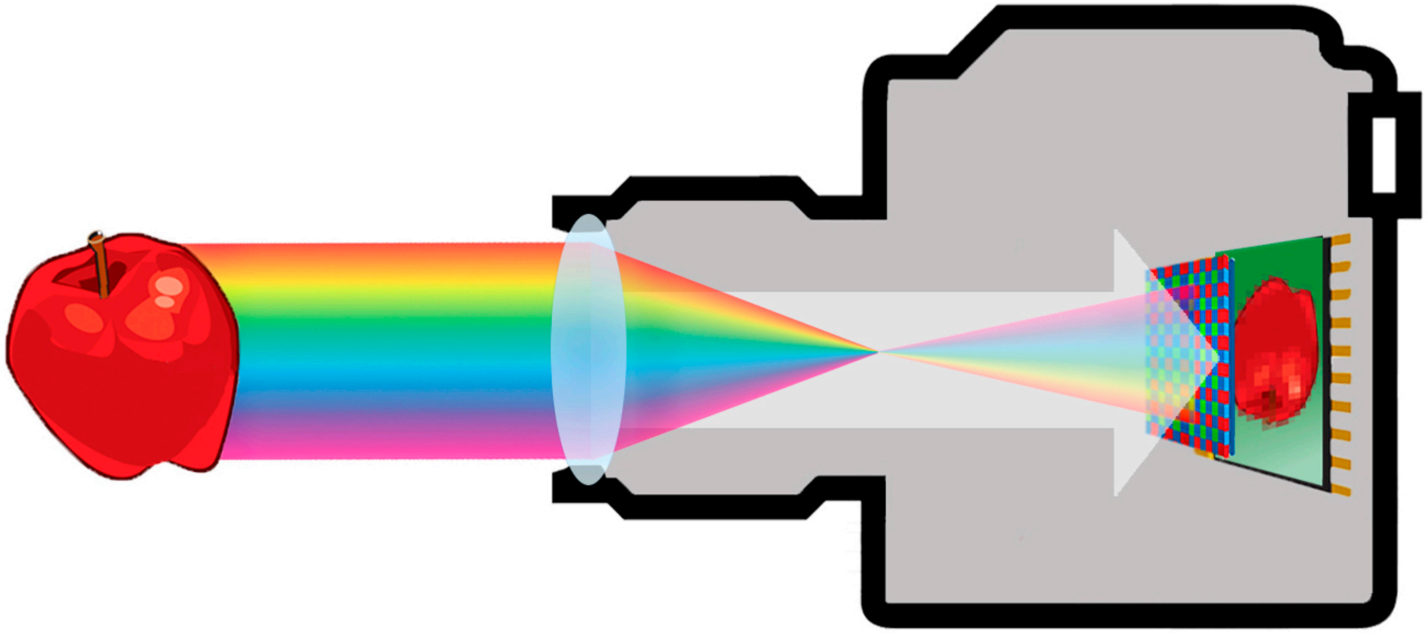
Mihály Bányai
banyai.mihaly@wigner.mta.hu
http://golab.wigner.mta.hu/people/mihaly-banyai/

- Formal languages
- Formalisation of knowledge as logic
- Dealing with uncertain knowledge

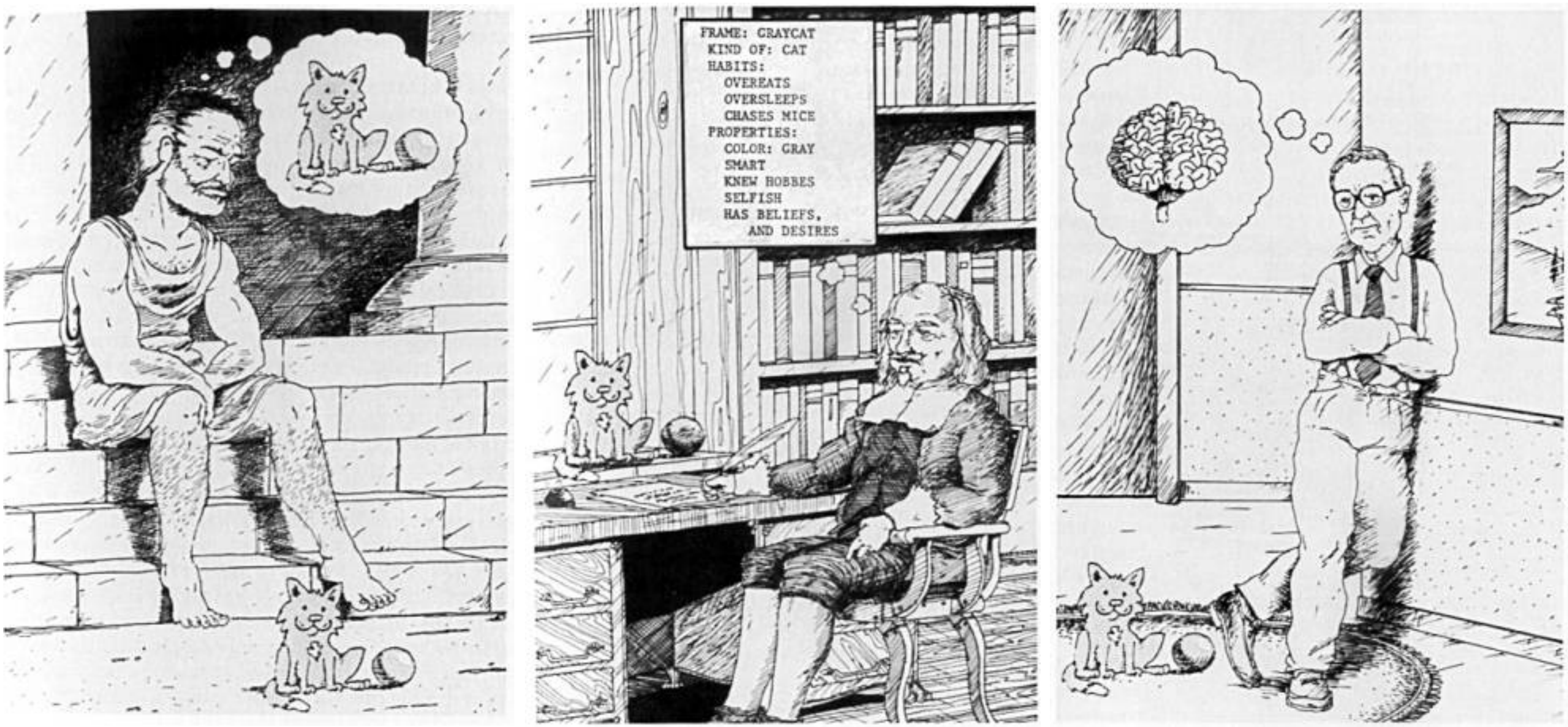
- Formal languages
- Formalisation of knowledge as logic
- Dealing with uncertain knowledge

What is a representation?

- the mapping without the mechanism, described in information-theoretic terms



Forms of representations



Cummins, 1989, Meaning and Mental Representation, MIT Press

Representation is not magical

- Consequently, you can always think in terms of it
  - Considering the statistical mapping from one system to another
  - Without necessarily considering the mechanisms giving rise to the mapping
- Examples
  - tachograph
  - cellular dynamics

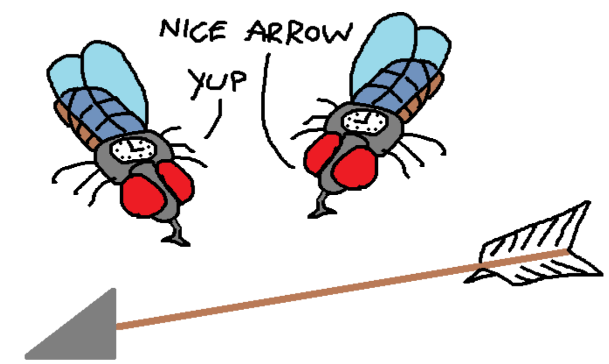
What do we need to handle knowledge?

- A language in which we can express the knowledge about the environment the brain has to handle
- the requirements of such a language will likely be different from those of a language used for conversations and literature
- we will use it to model the procedures with which the brain builds a model of its environment

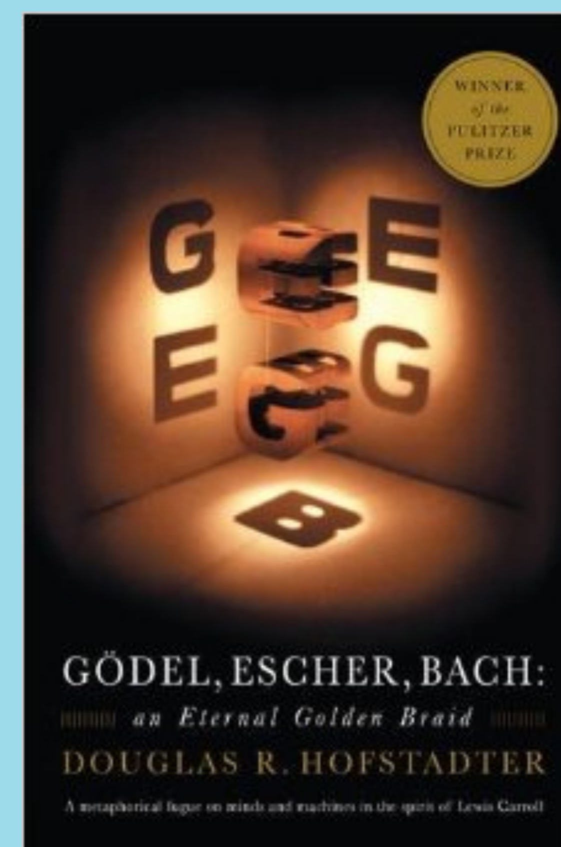
The need for formality

- animals and humans have to store information about how the world works
- this will be a set of propositions
- to handle such sets mathematically, we need a way to formalise propositions
- we could use natural languages like English
  - they contain all sorts of ambiguities
  - they are unnecessarily complex to model simple scenarios, as they address real-world context
    - we'd like to start modelling simple, experimentally well-testable problems
  - low-level perceptual phenomena, such as properties of a visual stimulus (textures, etc.) are cumbersome to describe

"Time flies like an arrow."
Groucho Marx



Pointer



Elements of formal languages

- we have a fixed set of symbols (words of the language)
- from these symbols we can assemble strings (sentences)
- we have a set of rules that decide if a string is a valid sentence (grammar)
  - grammatically correct sentences are called well-formed strings
- we have a set of sentences that we choose as axioms
  - the axioms define what your formal language can talk about. Symbols and grammatical rules are the form and axioms are the (first grain of) content
- formal languages with axioms are also called formal systems

Examples of formal languages

- geometry
  - symbols: points, lines and their relations
  - axioms: Euclid (or Bolyai-Gauss, etc.)
  - defined as a formal language by Alfred Tarski in 1959
- mathematics
  - symbols: sets, elements and relations
  - axioms: for example, the Zermelo-Fraenkel set theory with the axiom of choice (ZFC) - there are many others, depending on what kind of mathematical problems you want to handle
- programming languages
- music

Inference in formal languages

- Inference is the procedure with which we can produce new sentences from the axioms
- Axioms are the **knowledge base**, and with inference rules, we create new knowledge
- A formal language is completed by the set of rules that define how we may do so
- The strings that can be produced by the (repeated application of multiple) inference rules are called theorems
  - there are well-formed strings that are not theorems

Examples of formal inference

- mathematics
  - proofs
- programming languages
  - interpretation/compilation of the code

- Formal languages
- Formalisation of knowledge as logic
- Dealing with uncertain knowledge

What formal language should we use?

- There are infinitely many of them
- Some formal languages are explicitly designed to handle human knowledge in an intuitive way
- These are called **logics**, and their strings (sentences) **propositions**
- Symbols of logical languages always include certain relationship operators
  - NOT, AND, OR, denoted by $\neg, \wedge, \vee$
  - these allow for a special interpretation of the strings: they can be regarded as being true or false
- implication symbol: a shorthand notation $\neg A \vee B = A \rightarrow B$

Truth and falsity in logic

- true proposition -> theorem, given the axioms and the inference rules
- false proposition -> the negation (NOT + the proposition) is a theorem
- undecidable propositions -> not theorems, and their negation is also not a theorem

Logical languages

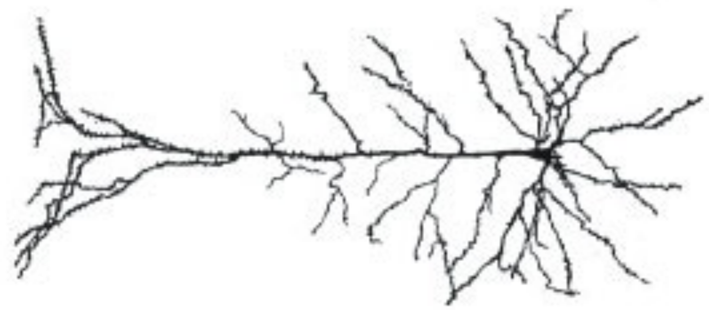
- Propositional logic: symbols represent whole natural language statements
    - X = “In summer it rains a lot” Y = “It is summer” Z = “It is raining”
    - sentences: $(X \wedge Y) \vee \neg Z$
- Predicate logic: statements can be formulated in a compositional way
  - symbols representing elements of a set of **objects** - e.g. animals
  - symbols representing properties and relationships of object - **predicates**
    - Predator(dog), Eats(fox,rabbit)
  - symbols representing **variables** that may stand for any object - x,y
  - symbols to tell wether we talk about all possible values of a variable or only one - **quantors:**
    - $\forall x$ - we state the following proposition for all possible values of the variable x
    - $\exists x$ - there exists at least one value of the variable x for which the proposition holds
  - sentences: $\forall x$ Predator(x) → Eats(x,rabbit)

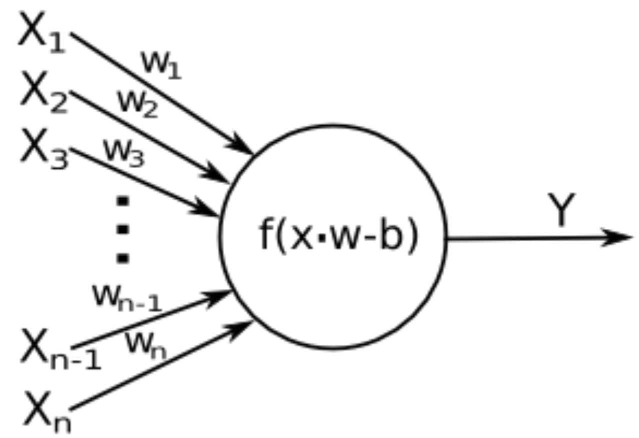
Inference in logic

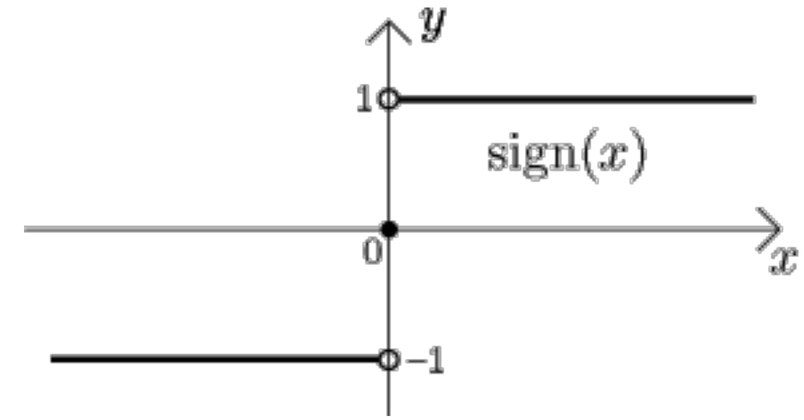
- In logical languages, inference rules can be defined as intuitive ways to find out whether a proposition is true
  - axiom set: AS = {All greeks wear togas. Socrates is Greek.}
  - proposition: X = {Socrates wears a toga.}
  - formalisation in predicate logic
    - basis set: humans
    - predicate symbols: Greek(x), WearsToga(x)
    - AS = $\forall$x Greek(x) -> WearsToga(x), Greek(Socrates)
    - X = WearsToga(Socrates)
    - Inference rule: if $\forall$x Pred1(x) -> Pred2(x) and Pred1(A) then Pred2(A)
- There are more complicated inference rules that can decide whether a proposition is a theorem more effectively

Pointer
Incompleteness theorems
http://www.scottaaronson.com/democritus/lec3.html

The neuron as a logical device

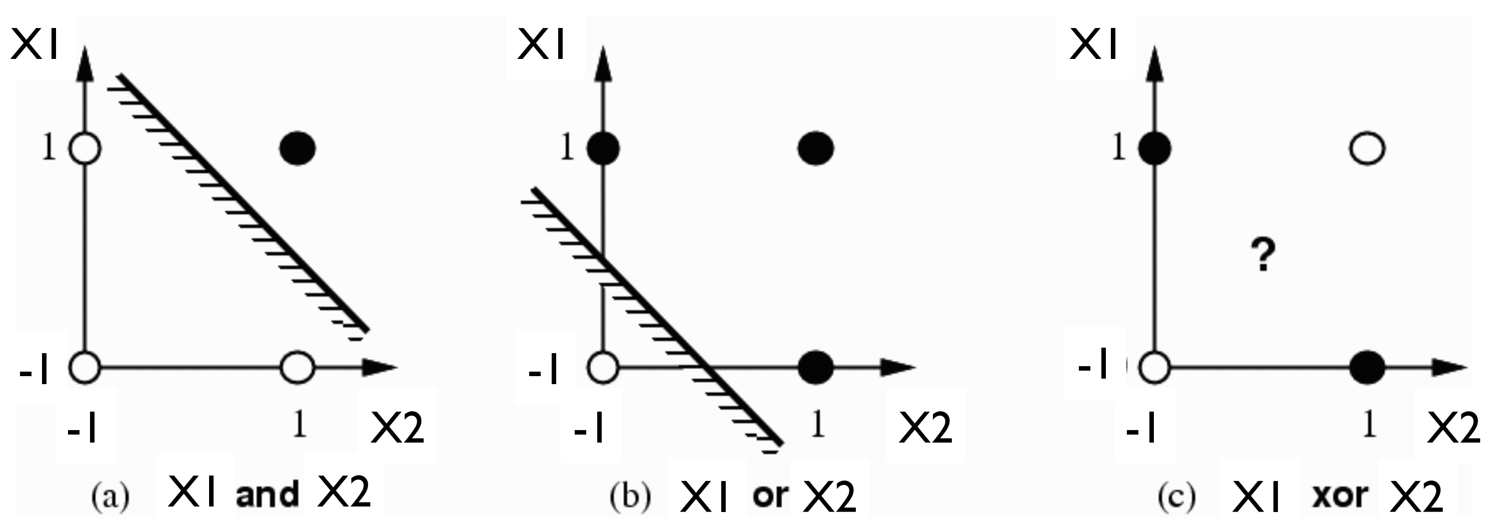






$$y = f(x_1 w_1 + x_2 w_2 - \theta)$$

$$y = 0 \longrightarrow \theta = x_1 w_1 + x_2 w_2 \longrightarrow x_2 = \frac{-w_1}{w_2} x_1 + \frac{\theta}{w_2}$$



A fairer treatment of undecidable sentences

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
    - this is a theorem
- what is our intuition about propositions like this?

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem
- what is our intuition about propositions like this?
  - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
    - this is a theorem
- what is our intuition about propositions like this?
    - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
    - both are undecidable

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem
- what is our intuition about propositions like this?
  - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
  - both are undecidable
  - but we see that if we new whether Achilles and Seamus were Greeks, the propositions would be decidable.

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem

- what is our intuition about propositions like this?
  - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
  - both are undecidable
  - but we see that if we new whether Achilles and Seamus were Greeks, the propositions would be decidable.
  - we don’t now this, but we might have additional knowledge about the world that we can include in the axiom set:

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
    - this is a theorem
- what is our intuition about propositions like this?
    - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
    - both are undecidable
    - but we see that if we new whether Achilles and Seamus were Greeks, the propositions would be decidable.
    - we don't now this, but we might have additional knowledge about the world that we can include in the axiom set:
        - AS' = {All greeks wear togas. Socrates is Greek. One out 600 people is Greek. 9 out of 10 people called Achilles are Greek. 1 out of 10000 people called Seamus is Greek. Only 1 out of 100 non-Greeks wears a toga.}

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem
- what is our intuition about propositions like this?
  - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
  - both are undecidable
  - but we see that if we new whether Achilles and Seamus were Greeks, the propositions would be decidable.
  - we don't now this, but we might have additional knowledge about the world that we can include in the axiom set:
    - AS' = {All greeks wear togas. Socrates is Greek. One out 600 people is Greek. 9 out of 10 people called Achilles are Greek. 1 out of 10000 people called Seamus is Greek. Only 1 out of 100 non-Greeks wears a toga.}
    - X and Y are just as undecidable as before

A fairer treatment of undecidable sentences

- AS = {All greeks wear togas. Socrates is Greek.}, X = {Socrates wears a toga.}
  - this is a theorem
- what is our intuition about propositions like this?
  - X = {Achilles wears a toga.} Y = {Seamus wears a toga.}
  - both are undecidable
  - but we see that if we new whether Achilles and Seamus were Greeks, the propositions would be decidable.
  - we don't now this, but we might have additional knowledge about the world that we can include in the axiom set:
    - AS' = {All greeks wear togas. Socrates is Greek. One out 600 people is Greek. 9 out of 10 people called Achilles are Greek. 1 out of 10000 people called Seamus is Greek. Only 1 out of 100 non-Greeks wears a toga.}
    - X and Y are just as undecidable as before
    - but we certainly have an idea about X being closer to a theorem than Y

- Formal languages
- Formalisation of knowledge as logic
- Dealing with uncertain knowledge

Compression of observations

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one
  - it would be too much data (even in hyperthymesia)

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one
    - it would be too much data (even in hyperthymesia)
    - it would be unnecessarily clumsy to access it

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one
  - it would be too much data (even in hyperthymesia)
  - it would be unnecessarily clumsy to access it
  - we couldn’t generalize - wouldn’t know what to expect when a brown dog shows up

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one
    - it would be too much data (even in hyperthymesia)
    - it would be unnecessarily clumsy to access it
    - we couldn’t generalize - wouldn’t know what to expect when a brown dog shows up
- I can compress well when I’m aware of typical regularities

given: ~100000 Byte

Compression of observations

- we cannot store every detail of all our memories - e.g. once I got bit by a white dog, once by a black one
    - it would be too much data (even in hyperthymesia)
    - it would be unnecessarily clumsy to access it
    - we couldn't generalize - wouldn't know what to expect when a brown dog shows up
- I can compress well when I'm aware of typical regularities

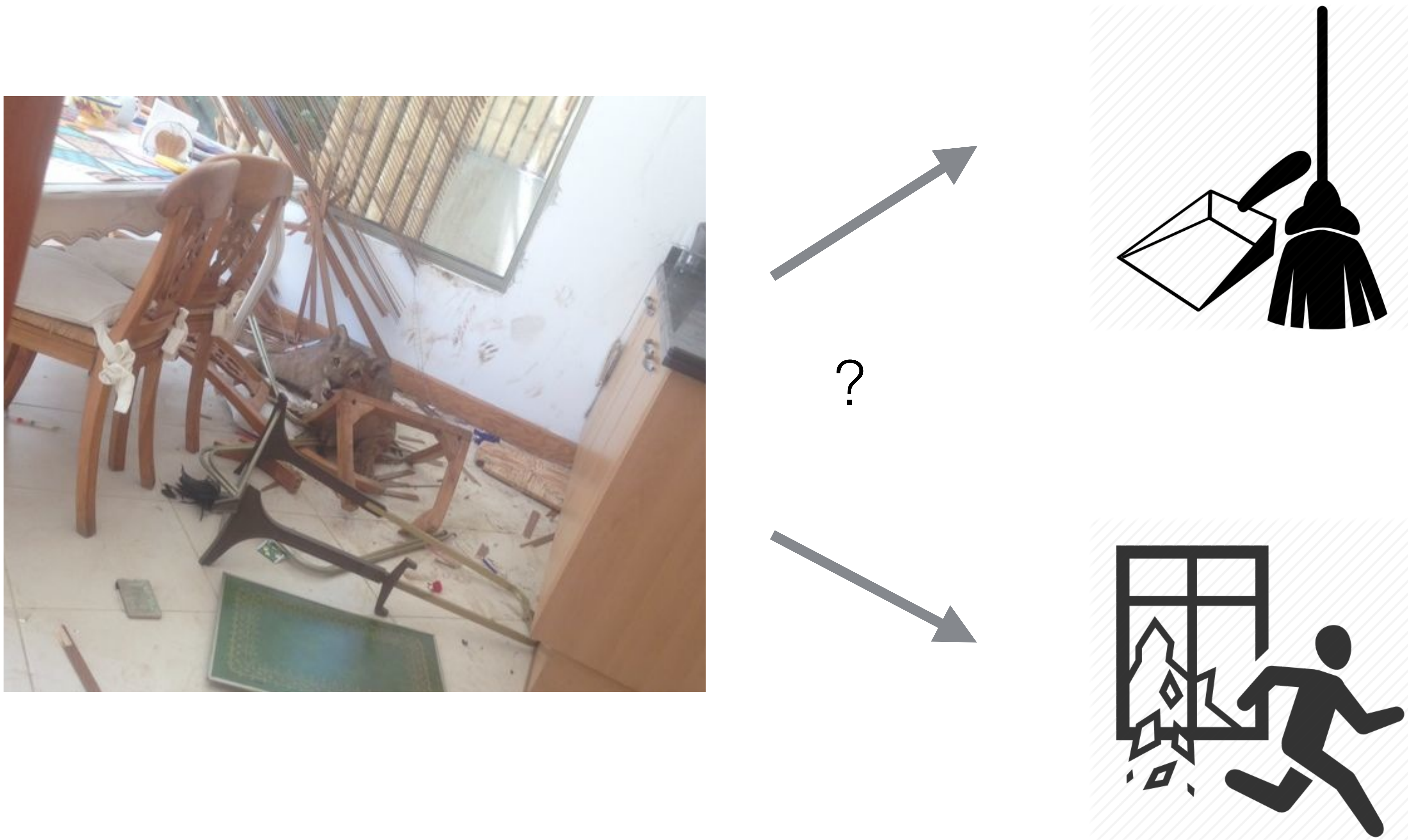
→ "untidy room with puma"

given: ~100000 Byte

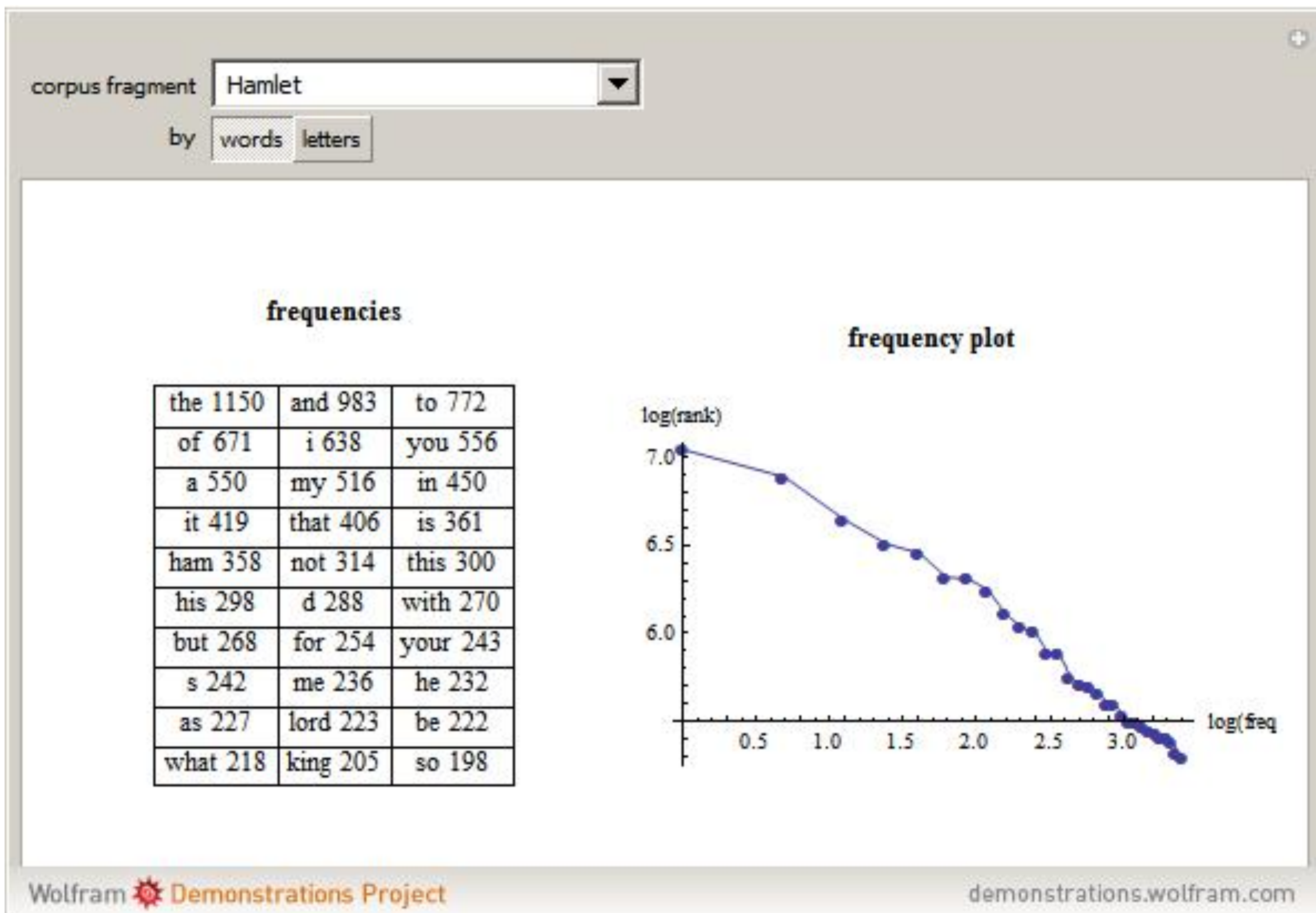
useful: ~40 Byte

Lossy compression

- Loosing information is a good thing



How to compress well?



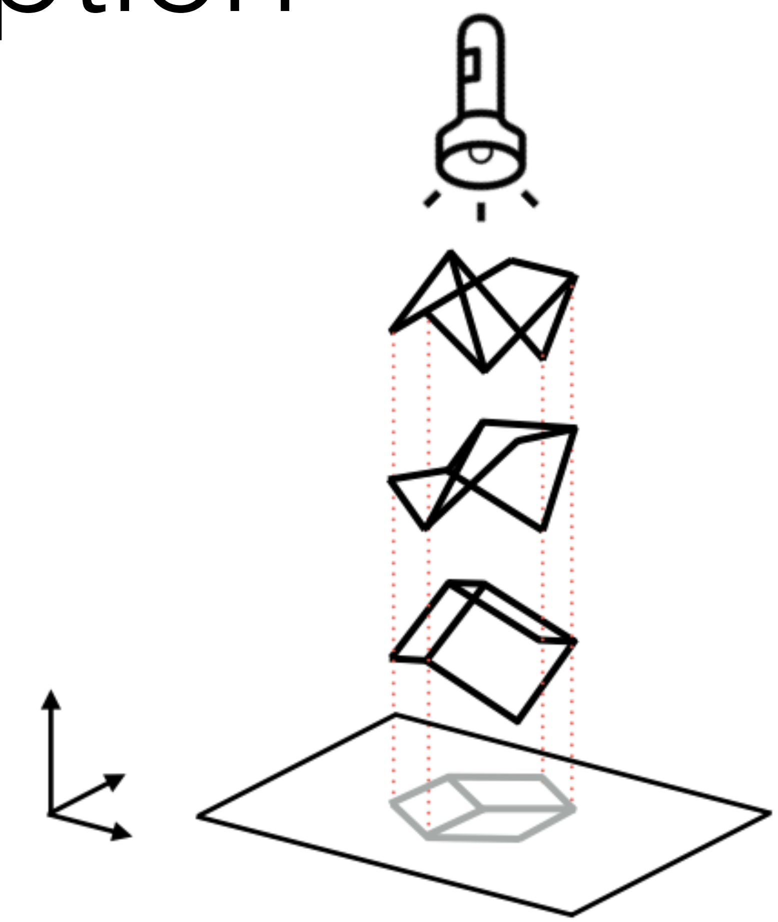
the 1150	and 983	to 772
of 671	i 638	you 556
a 550	my 516	in 450
it 419	that 406	is 361
ham 358	not 314	this 300
his 298	d 288	with 270
but 268	for 254	your 243
s 242	me 236	he 232
as 227	lord 223	be 222
what 218	king 205	so 198

- shorter descriptions should be used for more common cases
- to compress well, you have to know what is typical, and how likely different observations are

Why calculate with uncertainty?

- Why don't we just use the most likely value?

The need to handle uncertainty in perception



- perceptual indetermination is ubiquitous
- generalisability is key to function
    - if a brown dog bit me on Monday and a black dog bit me on Tuesday what will the spotty dog do on Wednesday?

Plausibility of a proposition

- In binary logic, a proposition is either a theorem of a certain axiom set, it contradicts it, or we can say nothing
- Let's extend this 3-valued evaluation into an infinite-valued one that can describe the plausibility of the proposition being a theorem by arbitrary precision

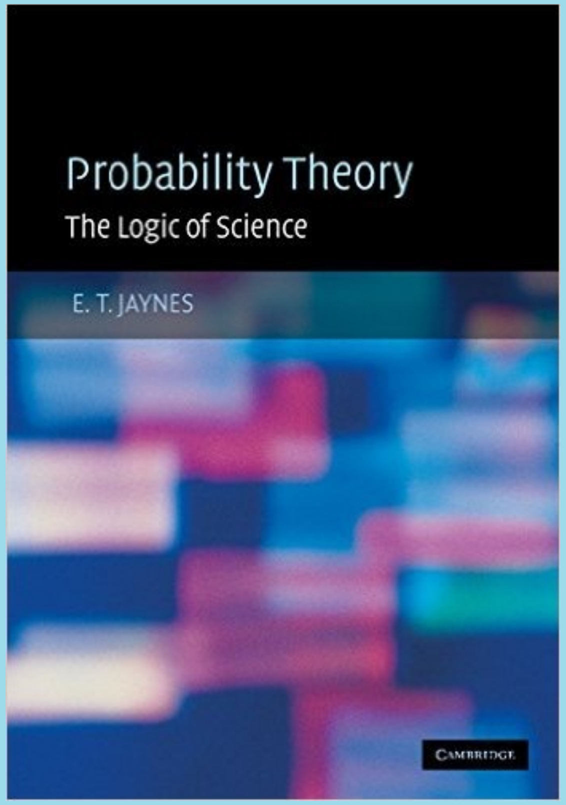
Numerical representation

- We can express any proposition as the assignment of a numerical value to a variable
  - "The river is 3.4 meters wide."
  - "This animal belongs to category 1."

What we want our plausibility measure to be like?

- the plausibility of a proposition X given an axiom set AS should be a real number, let's denote it by Pb(X | AS)
- consistency: starting from the same information (axioms), we should get the same plausibility value, no matter in what order we applied the inference rules (validly)
- the direction of change should be intuitive: if Pb(X | AS) increases, then Pb(X∧Y | AS) should also increase, and Pb(¬X | AS) should decrease
- Cox theorem says that if these are fulfilled, we obtain probability calculus for the description of plausibilities: Pb = Pr
    - (we need slightly more precise versions of the requirements for this to be technically true, but the basic idea is the same)

Pointer



Probability calculus

- we decide that the probability of a theorem (certainly true proposition) is 1. We don't lose any expressive power doing this.

- consequence: the probability of all mutually exclusive propositions sum up to 1

$$Pr(X \mid AS) + Pr(\neg X \mid AS) = 1$$

- we say that we are looking for the probability of X **conditioned** on AS

- we have two inference rules to derive the probabilities of propositions using the already known ones

Product rule

$$Pr(X \wedge Y \mid AS) = Pr(X \mid Y \wedge AS)Pr(Y \mid AS)$$

- what's the probability of Bill watching a football game at any time?
    - there's a 0.3 probability of a game going on
    - if there's a game, the probability of Bill watching it is 0.7
    - the answer is 0.21
- a direct consequence of this rule is the definition of conditional probability

$$Pr(X \mid Y \wedge AS) = \frac{Pr(X \wedge Y \mid AS)}{Pr(Y \mid AS)}$$

Sum rule

$$\overset{X}{Pr(x = 1 \mid AS)} = \sum_{i=1}^{N} Pr(\overset{X}{x = 1} \wedge \overset{Y_i}{y = i} \mid AS)$$

- x - rain, y - night or day
- let’s say that the probability of raining at night is 0.3, at daylight 0.2
- let’s say the night lasts for 10 hours - Pr(y=night) ~ 0.4
- according to the product rule:
  - Pr(y=night,x=rain) = 0.3 x 0.4, Pr(y=day,x=rain) = 0.2 x 0.6
- according to the sum rule, the probability of rain regardless of the time of the day is 0.24
- also called marginalisation

Bayes theorem

- another direct consequence of the product rule

$$Pr(X \mid Y \wedge AS) = \frac{Pr(Y \mid X \wedge AS)Pr(X \mid AS)}{Pr(Y \mid AS)}$$

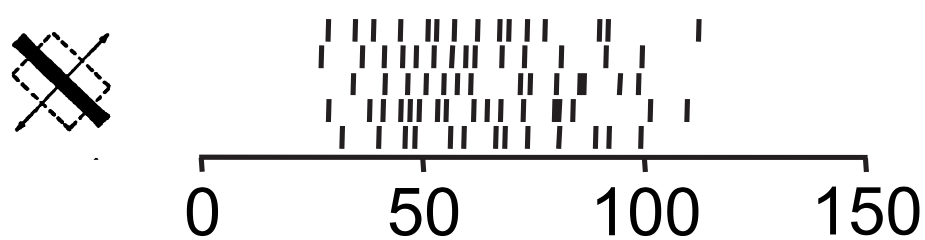
  - X - someone has TB
  - Y - a test for TB gives a positive result
  - we know that the test gives a positive IF the patient has TB with 0.9 probability
  - what is the probability of someone has TB IF the test came out positive?
  - have to take into account base rates - how probable a priori is it for someone to have TB, and how probable is it for the test to give a positive in any condition
- a Bayesian is someone or something using probability theory - no more, no less

Notational simplicity

- that makes things more complicated
- we often leave conditions implicit
- Pr(X) means Pr(X | AS), where AS is the axiom set (knowledge base), all the information that was taken into account when quantifying the probability of X
- as the knowledge base is always in the condition of all probabilities related to a given problem, this omission does not cause any technical problem
- but we shouldn't forget that it's always there

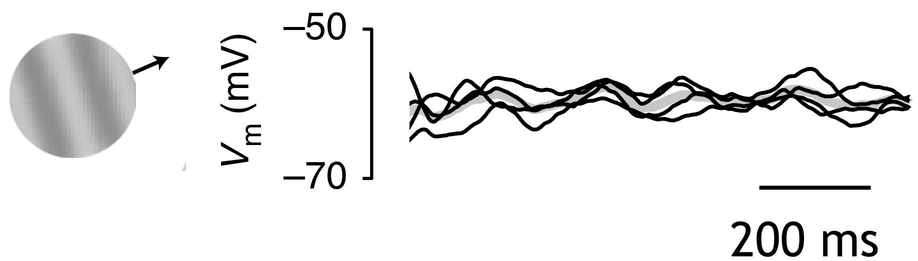
Variability in the neural responses

V1 spike trains



Gur & Snodderly, Cereb Cortex 2006

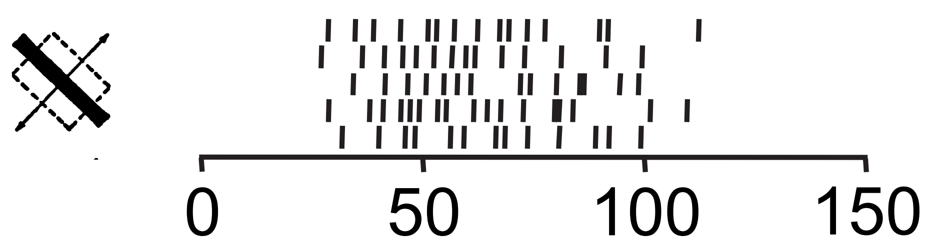
V1 membrane potentials



Finn et al, Neuron 2007; Churchland et al, Nat Neurosci 2010

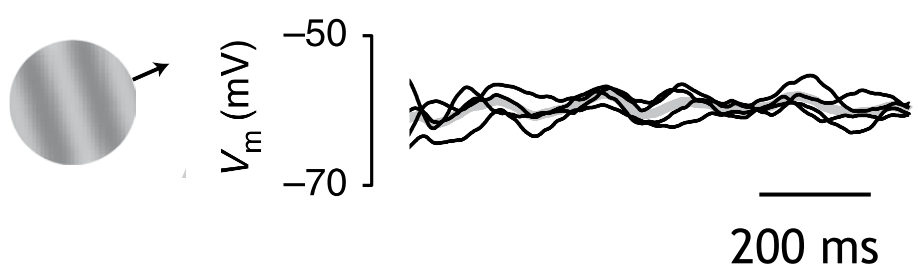
Variability in the neural responses

V1 spike trains



Gur & Snodderly, Cereb Cortex 2006

V1 membrane potentials

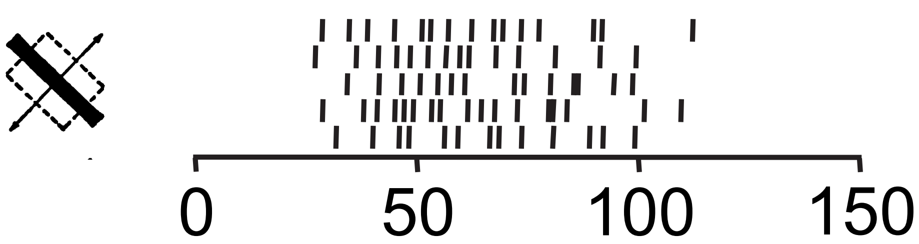


Finn et al, Neuron 2007; Churchland et al, Nat Neurosci 2010

- The fact that neurons respond differently to the same stimulus gives a hint about the nervous system handling uncertainty

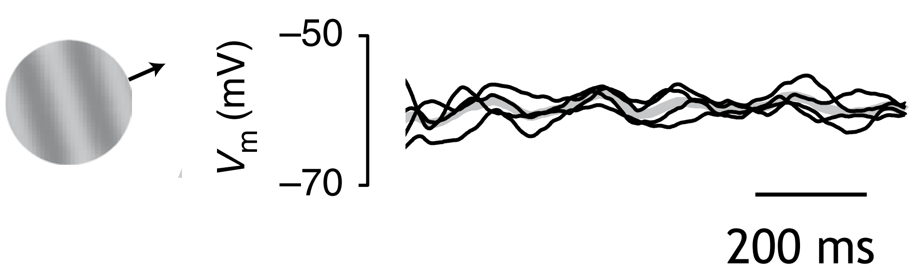
Variability in the neural responses

V1 spike trains



Gur & Snodderly, Cereb Cortex 2006

V1 membrane potentials

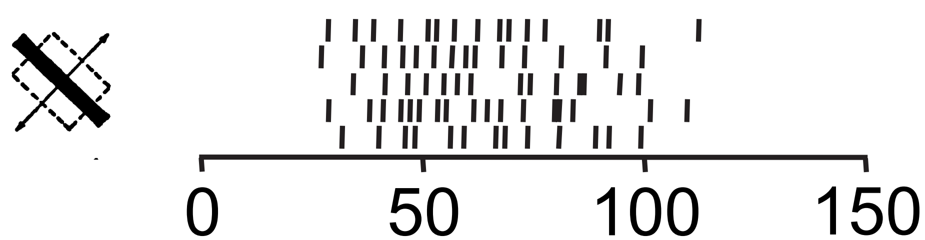


Finn et al, Neuron 2007; Churchland et al, Nat Neurosci 2010

- The fact that neurons respond differently to the same stimulus gives a hint about the nervous system handling uncertainty
- By averaging these responses, we get the RFs

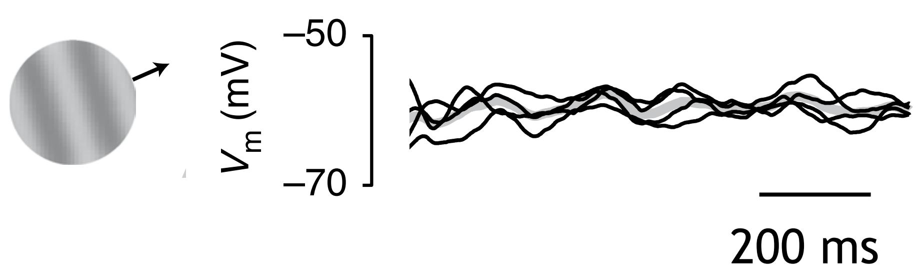
Variability in the neural responses

V1 spike trains



Gur & Snodderly, Cereb Cortex 2006

V1 membrane potentials

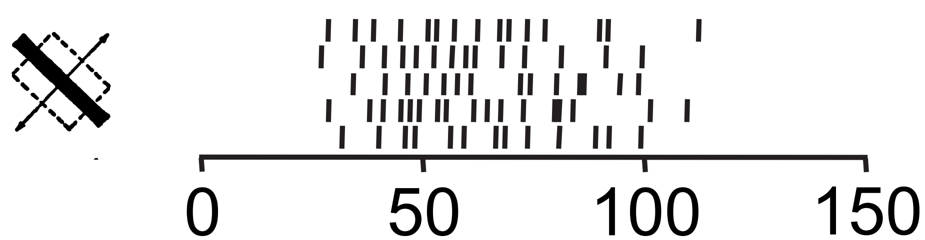


Finn et al, Neuron 2007; Churchland et al, Nat Neurosci 2010

- The fact that neurons respond differently to the same stimulus gives a hint about the nervous system handling uncertainty
- By averaging these responses, we get the RFs
- But when we are trying to figure out cortical computations, we can postulate that this variability serves a purpose.

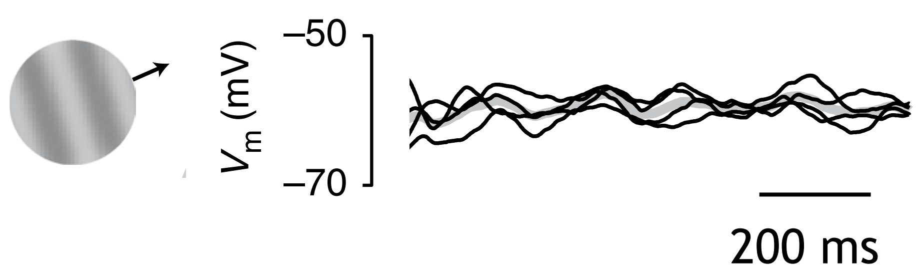
Variability in the neural responses

V1 spike trains



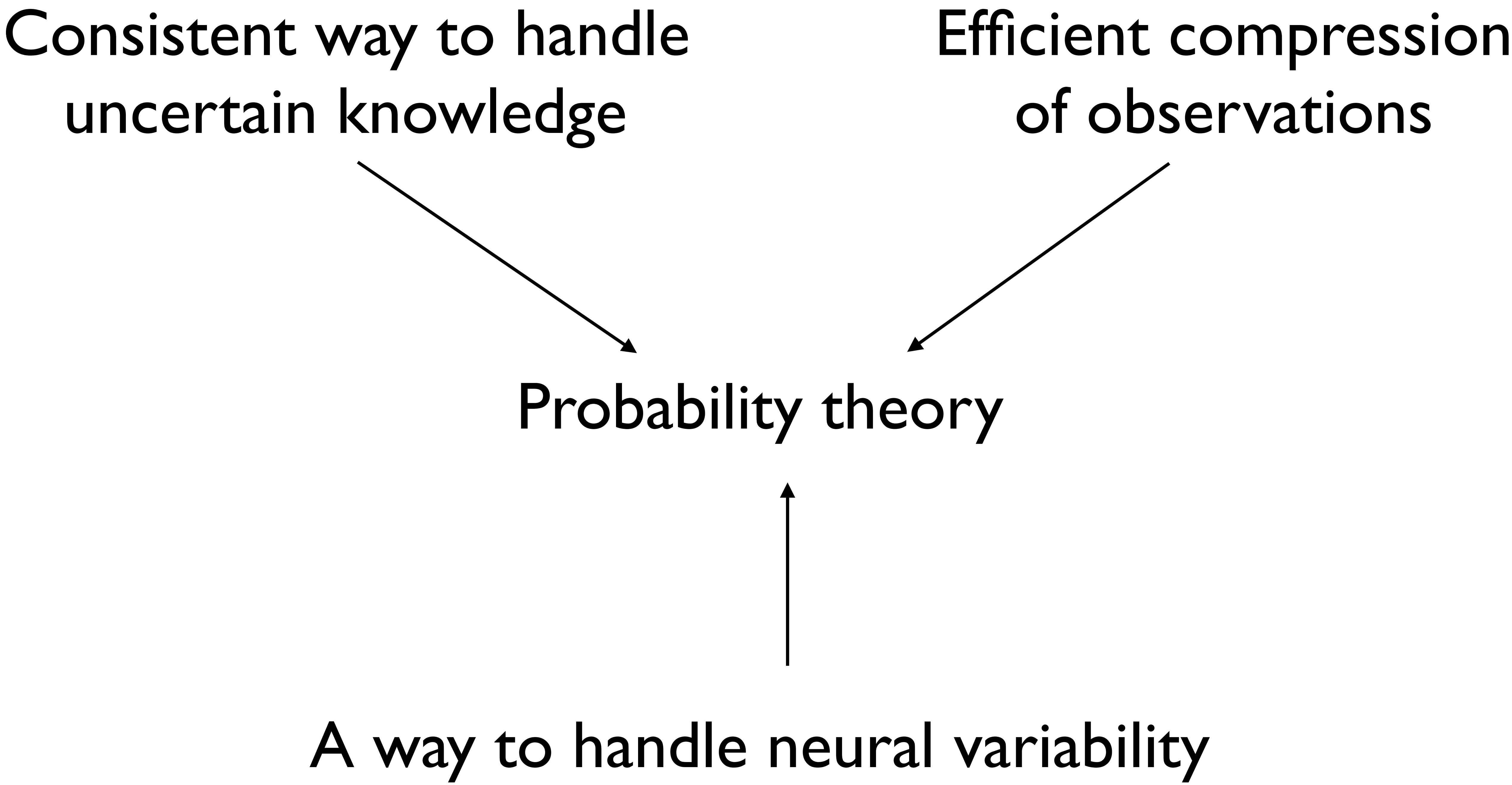
Gur & Snodderly, Cereb Cortex 2006

V1 membrane potentials



Finn et al, Neuron 2007; Churchland et al, Nat Neurosci 2010

- The fact that neurons respond differently to the same stimulus gives a hint about the nervous system handling uncertainty
- By averaging these responses, we get the RFs
- But when we are trying to figure out cortical computations, we can postulate that this variability serves a purpose.
- We can try to predict this variability assuming that the brain conducts probabilistic inference.


Consistent way to handle uncertain knowledge
Efficient compression of observations
Probability theory
A way to handle neural variability

Sidenote - interpretations of probability

Pointer
Kolmogorov axioms

- Frequentist
  - probability can be interpreted in repeated experiments as the relative frequency of an outcome among all trials
- Information-based (Bayesian, Laplacian)
  - probability describes the uncertainty of the information an observer has about some phenomenon
  - Subjective (de Finetti)
    - probability represents personal beliefs
  - Logical (objective, Jaynes)

Pointer
How quantum mechanics relates to probability theory?
http://www.scottaaronson.com/democritus/lec9.html

Sidenote - other attempts to quantify uncertainty

- Null hypothesis significance testing
  - a heuristic to assess the plausibility of a proposition using some elements of probability theory
  - you can do them by pencil and paper if needed
- Fuzzy logic
- According to the Cox theorem, these either end up with the same plausibilities as probability theory, or they become inconsistent at some point

The way forward

- Now we have a framework of handling knowledge that we introduced as a natural extension of logic to uncertain cases
    - coincidentally, this happens to be probability calculus, for which there is a vast amount of techniques readily available
- We have to develop tools to formalise real problems of perception (representation, inference and learning) in terms of probability theory
- Then we can move on to make predictions about behaviour and ultimately neural activity