Cognitive science for machine learning 1:
What is cognitive science?

Nick Chater

OVERVIEW

1. Cognitive science meets machine learning
2. Historical background: introspection, behaviourism, the computer metaphor
3. Levels of explanation

1. COGNITIVE SCIENCE MEETS MACHINE LEARNING

COGNITIVE SCIENCE AS REVERSE ENGINEERING

- Much of cognition involves induction: finding patterns in data (Josh, next session)
- Machine learning provides an engineering toolkit for inductive inference
  - Cognitive science draws on machine learning for hypotheses
  - And provides machine learning with interesting challenges

But, in practice, the concerns of cog sci and ML may appear to clash

Cognitive science

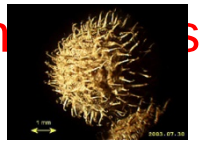
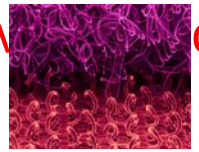
- Focus on specific experimental paradigms
- Embedded in psychology in linguistics
- Aiming to be “cognitively” and/or “neurally” plausible
  - Do we know enough to impose such constraints?

Machine learning

- Focus on standard learning problems
- Squeezing out an extra 1%
- Embedded in computer science and engineering
- Aim for a working system, whether mimicking the brain or not

COGNITIVE SCIENCE AND MACHINE LEARNING: WHY ARE WE ALL HERE?

- The biggest single empirical constraint on any model of a cognitive process is that it solves the task successfully
- And without a powerful computational tools, it is impossible even to understand the problems the brain solves
  - Cognitive science needs machine learning
- The *only* working example of an intelligent learning system is the brain
- And many machine learning problems are defined in terms of human cognition (from object recognition, to machine translation)
  - Machine learnin[illegible]s cognitiv[illegible]ce (cf biomimetics)

burrs   velcro

2. HISTORICAL BACKGROUND:

INTROSPECTION, BEHAVIORISM, THE COMPUTER METAPHOR

WUNDT AND THE BEGINNING OF EXPERIMENTAL PSYCHOLOGY

- First experimental psychology laboratory at Leipzig, in 1879

- **Structuralism**: "Human mental experience, no matter how complex, can be viewed as blends or combinations of simple processes or elements."

  - Influenced by J.S. Mill's – **mental chemistry**.

  - But rather than **computational components**, building blocks are **subjective experience** (qualia)

Wilhelm Wundt (1832–1920). *[Archives of the History of American Psychology, University of Akron].*

THE INTROSPECTIVE METHOD

- **Experimental Psychology** "…the investigation of conscious processes in the modes of connection peculiar to them"

- **Method** – Systematic introspection, under experimental control

THE METHOD IN ACTION

- very simple stimuli → verbal report

O listens to a metronome. After a time the beats form rhythmic groupings and various conscious experiences may be reported, such as, at the end of a group there is an impression of an "agreeable whole". He then tries to describe the qualities of this experience, such as feelings of pleasure or displeasure, tension or excitement

- **Attempt** to isolate the "elements of consciousness" out of
  which more complex mental events are made.

- **Metronome**: single beat = a sensation
  Combination into rhythms = an idea.

THE COLLAPSE OF INTROSPECTION

- the imageless thought controversy: Wundt vs Külpe
- Different labs produced very different results
  e.g., Leipzig vs. Cornell
- Introspection can change the phenomenon observed

THE UNRELIABILITY OF INTROSPECTION IN PERCEPTION

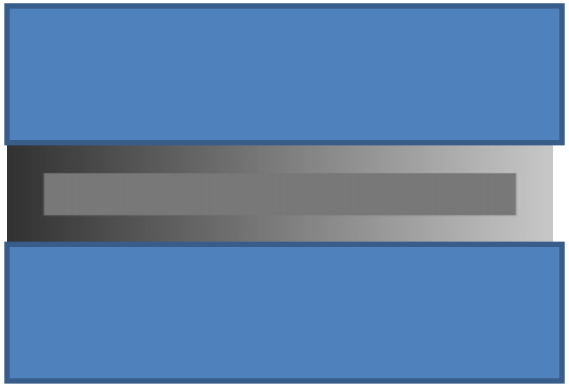
- How much do can we introspect about the retinal image? (demo)

- Endless visual illusions

THE UNRELIABILITY OF INTROSPECTION IN PERCEPTION

- How much do can we introspect about the retinal image? (demo)

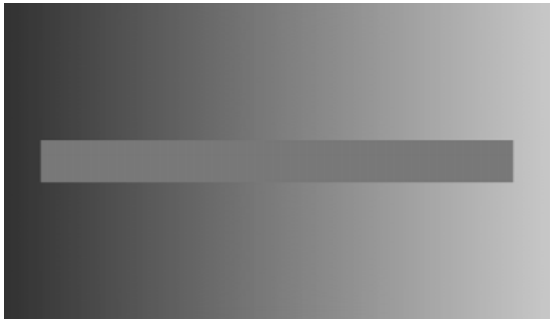
- Endless visual illusions



THE UNRELIABILITY OF INTROSPECTION IN PERCEPTION

- How much do can we introspect about the retinal image? (demo)

- Endless visual illusions

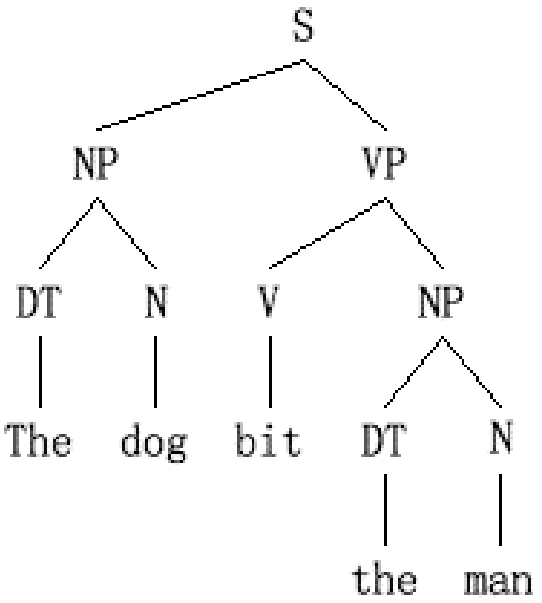


LITTLE BETTER FOR KNOWLEDGE OR DECISION

- In AI, experts notably can't articulate what they know
- And we can't articulate our "naive" theories
  - why Lenat's CYC is difficult
  - why linguistics is difficult
- People often can't explain their behaviour (Johansson & Hall's choice blindness, e.g., Science 2005)

INTROSPECTION IS HUGELY POWERFUL BUT REPORTS OUTPUT, NOT PROCESS

- High-level vision
- Action intentions
- Linguistic intuitions
- Mathematical intuitions
- Concepts
- Causality...



NB. THE SPECIFICATION OF ML PROBLEMS; AND THE MEANS OF SOLVING THEM, DRAW ON INTROSPECTION

BEHAVIORISM

- Perhaps psychology is not about subjective experience but objective behaviour

- Emergence of behaviorism: very strong claim that psychological laws should be framed over direct relationships between physically characterised stimuli and/or responses

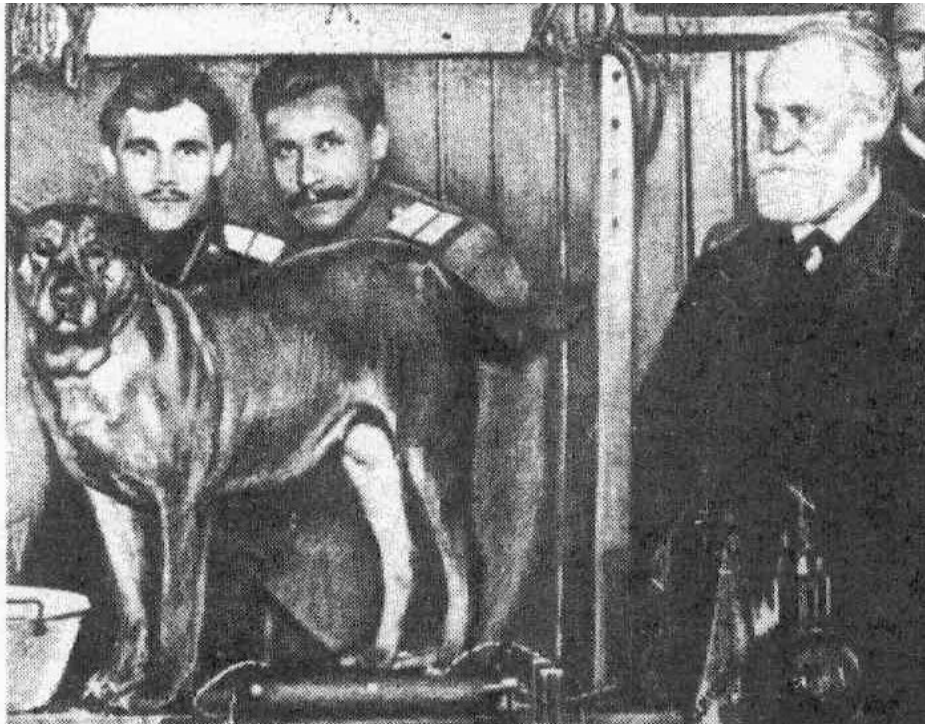
- Main method: looking at animal learning: where stimuli and learning can be carefully controlled and measured

SETTING THE STAGE: PAVLOV (1849 – 1936)

Physiologist studying dog digestion, found, by chance, that dog salivation was triggered by a bell that usually preceded food

Thus, the dog had learned an CS-US (bell-food) association

Classical or Pavlovian conditioning



J.B. WATSON (1878 – 1959)
FOUNDER OF BEHAVIORISM

- **Watson** dismissed introspection as hopelessly unscientific. To be replaced with:
  - **Study** of observable behaviour
    (no need to look inside the 'black box')
  - **Explained** via S-S and S-R links
  - **Thought** is movement of the larynx (hidden behaviour)
  - **Using** careful 'operational' definitions of all psychological terms (or tried to)

John Broadus Watson (1878–1958). *[Archives of the History of American Psychology, University of Akron].*

B. F. SKINNER (1904 – 1990)
RADICAL BEHAVIORISM

Further developed behavorist research, working on operant or Skinnerian conditioning (building relationships between Responses and Rewards)

Produced a vast research programme on learning in pigeons in the “Skinner box”

Schedules of food reward determined by, e.g., lever pressing

Starting point for ideas of reinforcement learning in neuroscience and machine learning

B. F. Skinner with one of his subjects. *[Ken Heyman].*

AND IT *WAS* RADICAL

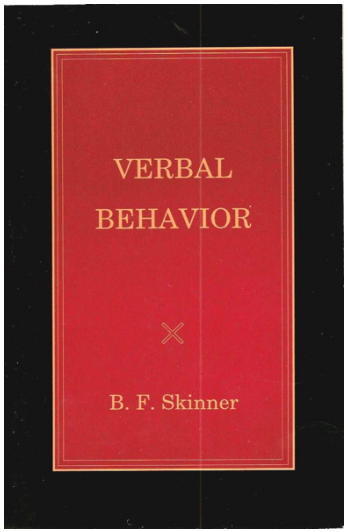
Aimed to explain *all* human behaviour, including language

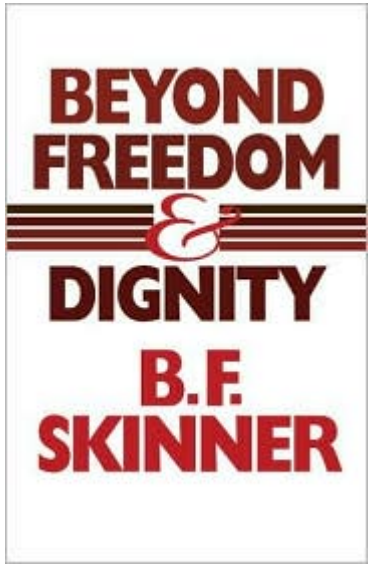
No real theory of perception or motor control

Typically no attempt to link with the brain

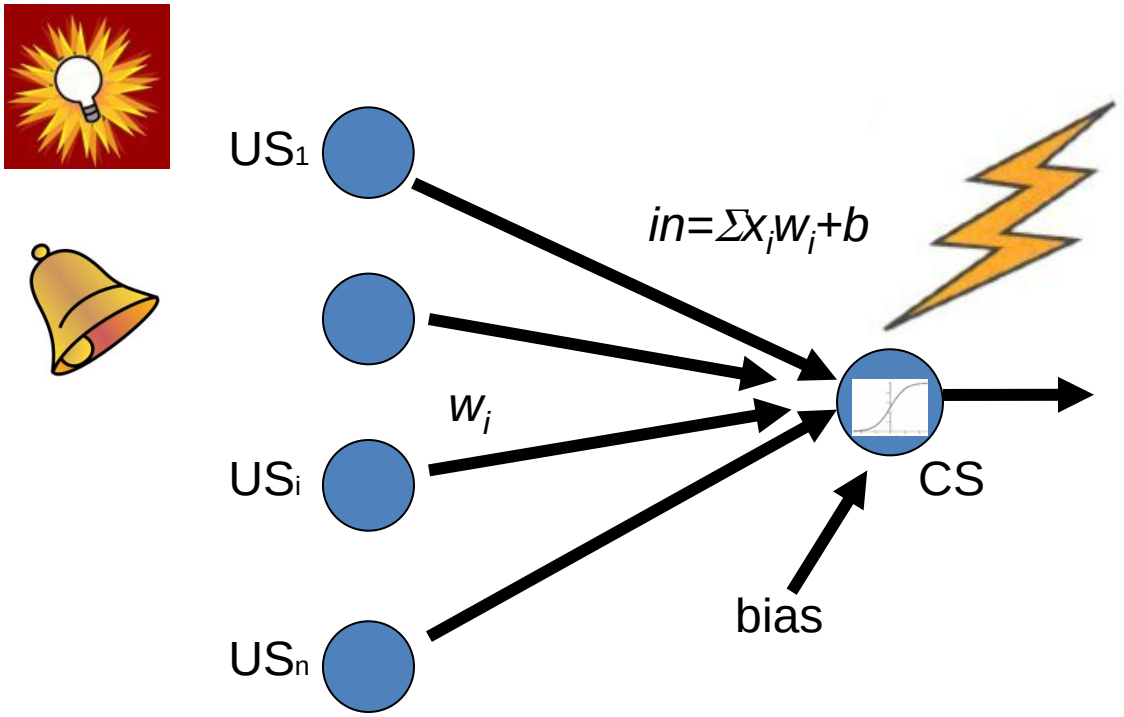
Reinforcement history explicitly viewed as the correct alternative to our view of ourselves as reasoning beings.

Only innate structure is principles of association



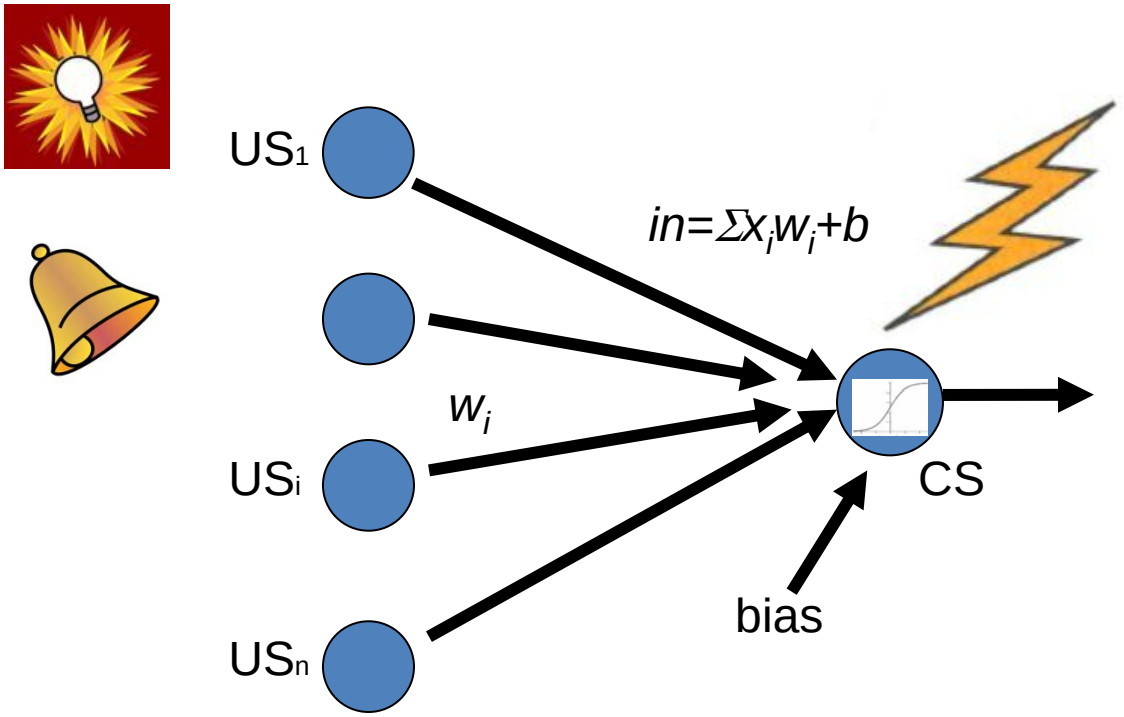


SOMEWHAT ANACHRONISTICALLY BEHAVIORISM IN COMPUTATIONAL TERMS



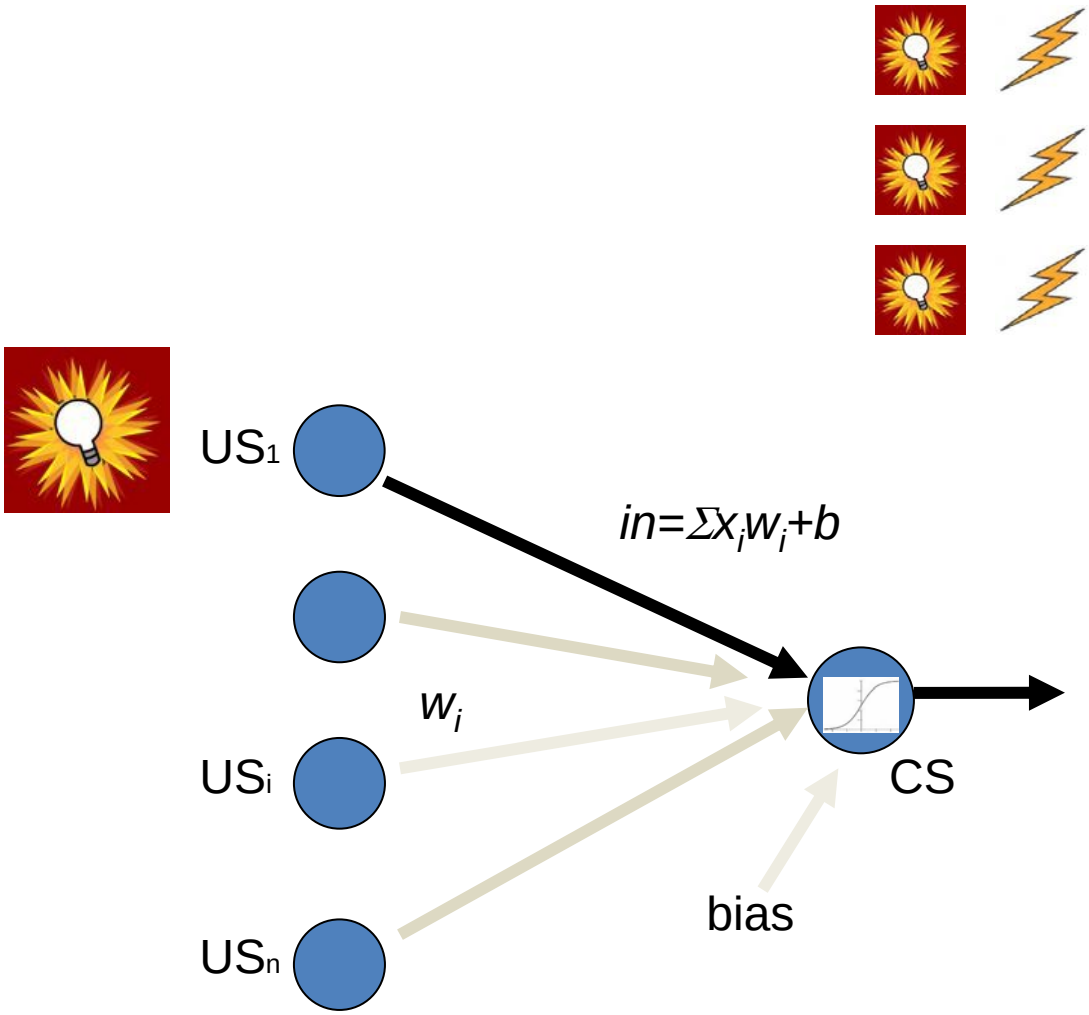
Hebbian learning: associate events near in time and space

BY CONTRAST RESCORLA-WAGNER (1970) MODEL OF CLASSICAL CONDITIONING

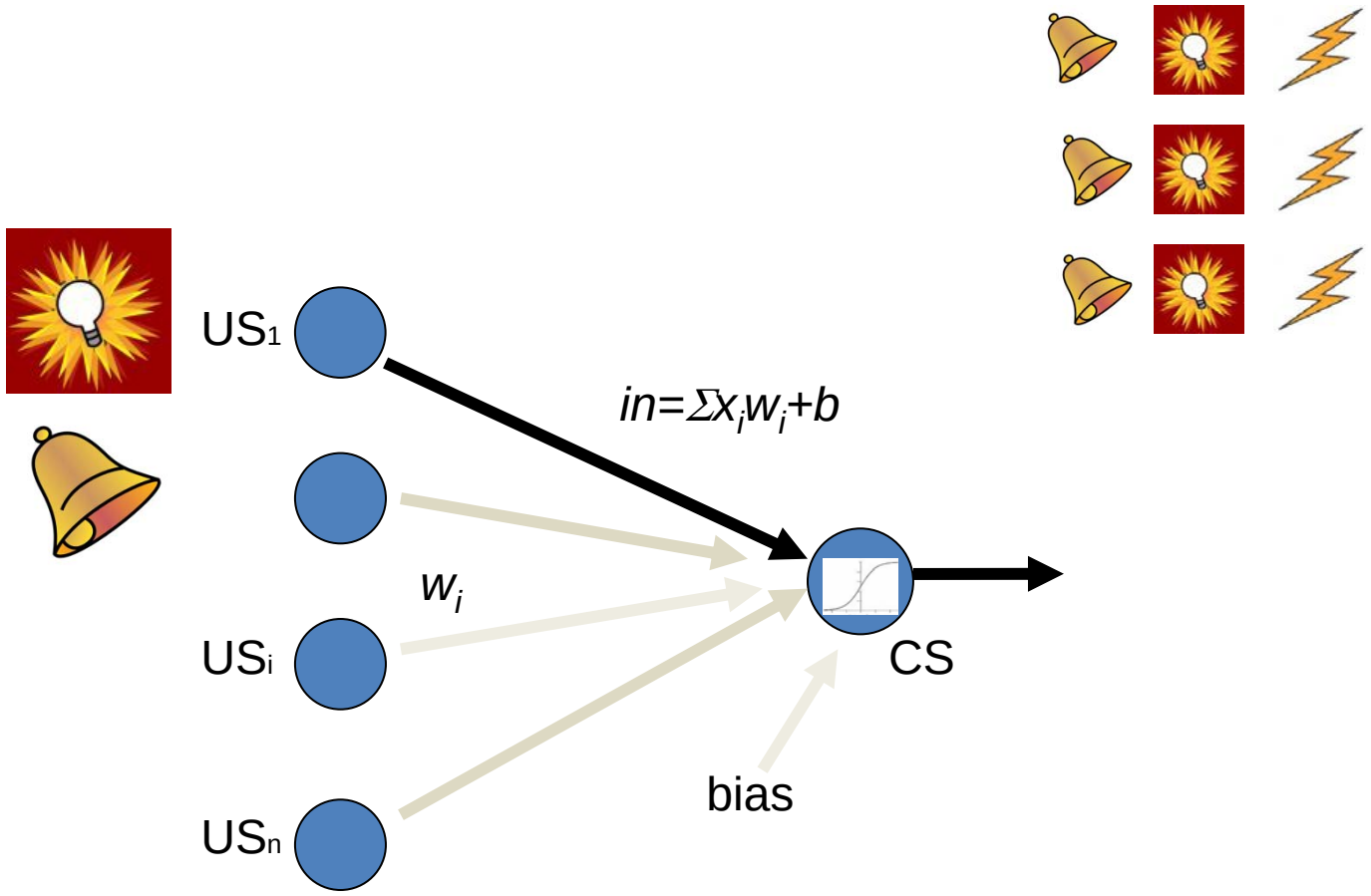


Error-driven learning: learn only form mistakes;
Crucial evidence: Kamin blocking (1968)

KAMIN BLOCKING : TRAINING PHASE 1



KAMIN BLOCKING: TRAINING PHASE 2



NO ERROR, SO NO FURTHER LEARNING

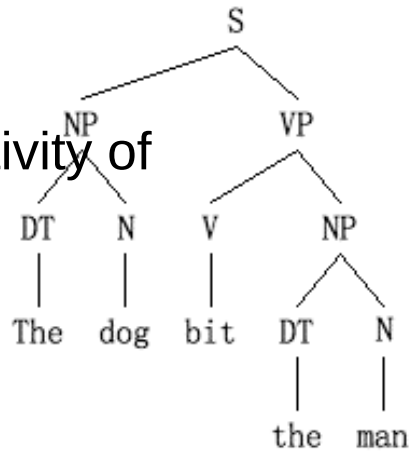
PROBLEMS FOR BEHAVIORISM

- Perception and motor control: Associations between categories “lever,” “press,” - but this is circular – these categories must be explained.

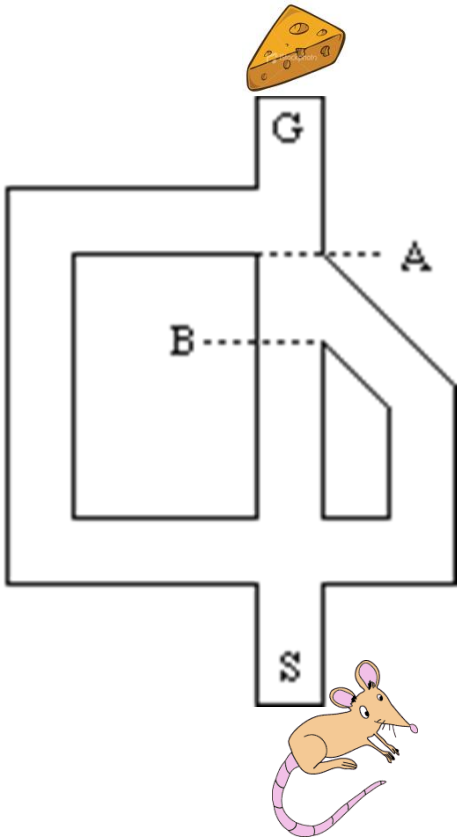
  (nb. behaviorists could say nothing useful about behavior in terms of physical categories – mass, force, brightness)

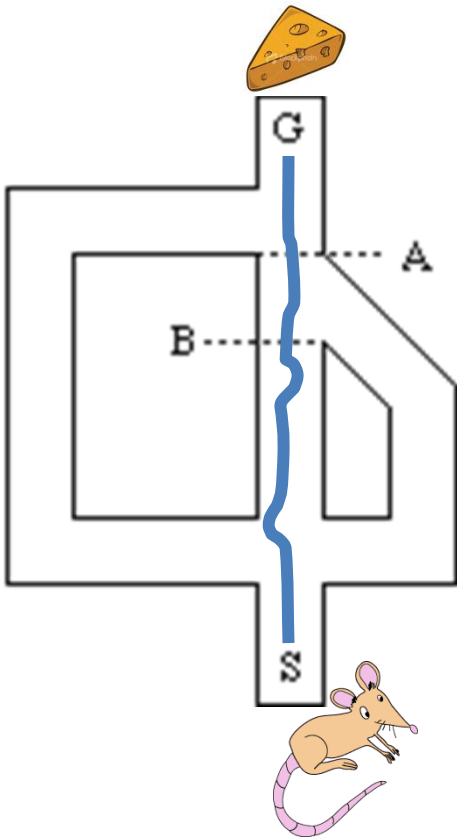
  So: <u>circularity</u>

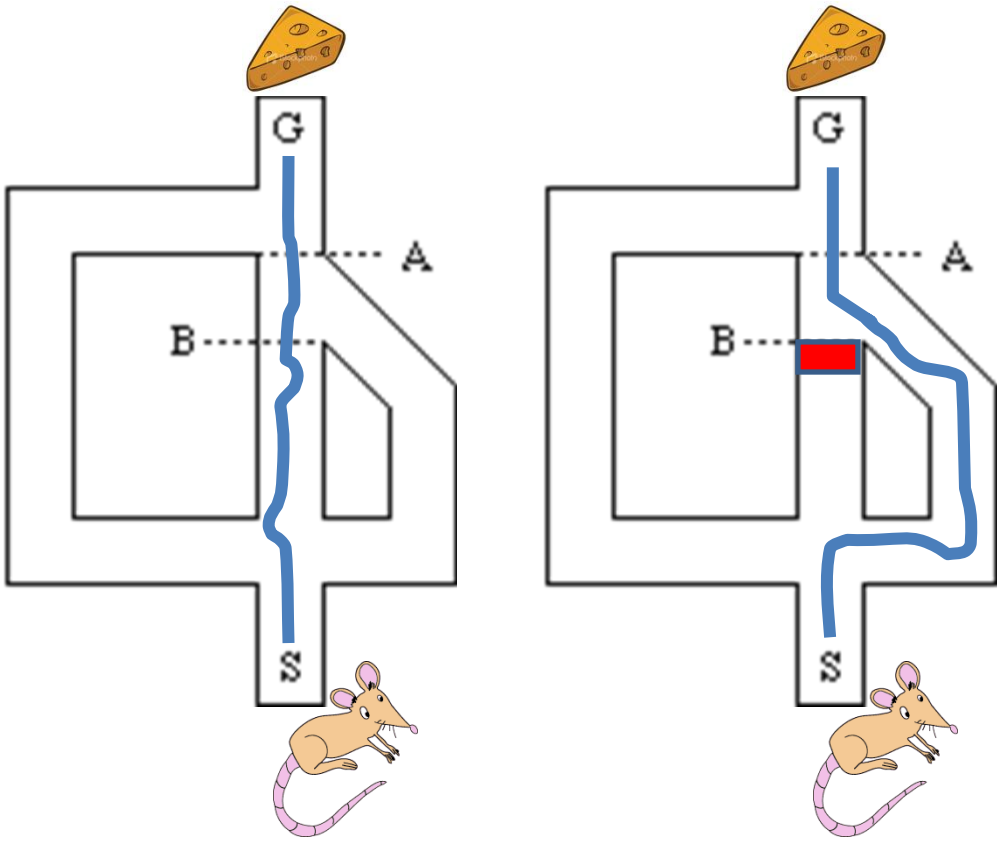
- Language : Chomsky showed that the infinite creativity of language cannot be explained in S-R terms
  (we learn rules for language, not S-R associations)

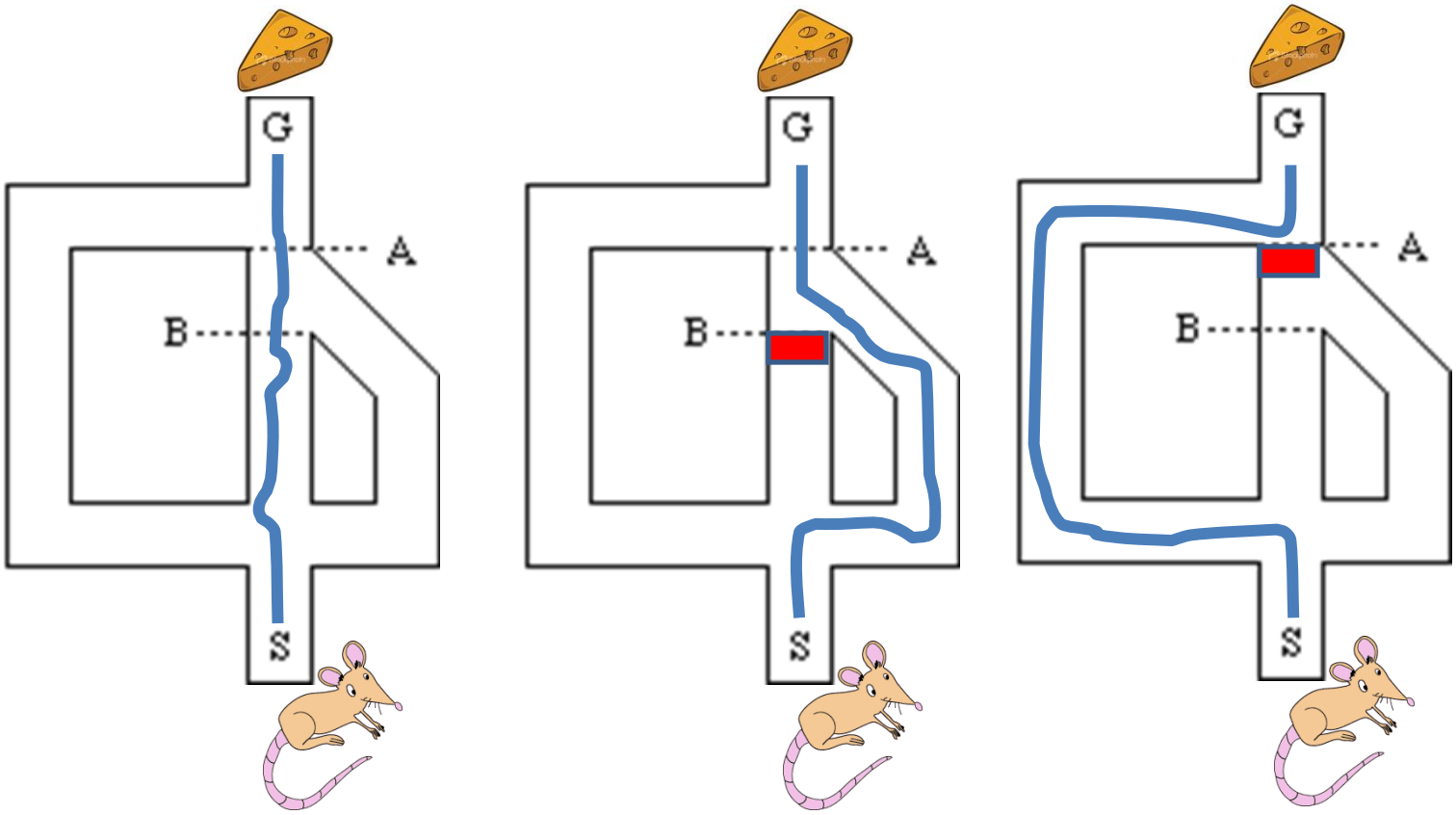


- **Flexibility of behaviour** i.e., behaviour guided by reasoning to the solution of a novel problem
- **Example** Rats tend to take the shortest route through a maze rather than the one that has been most reinforced **(Hull and Tolman)**

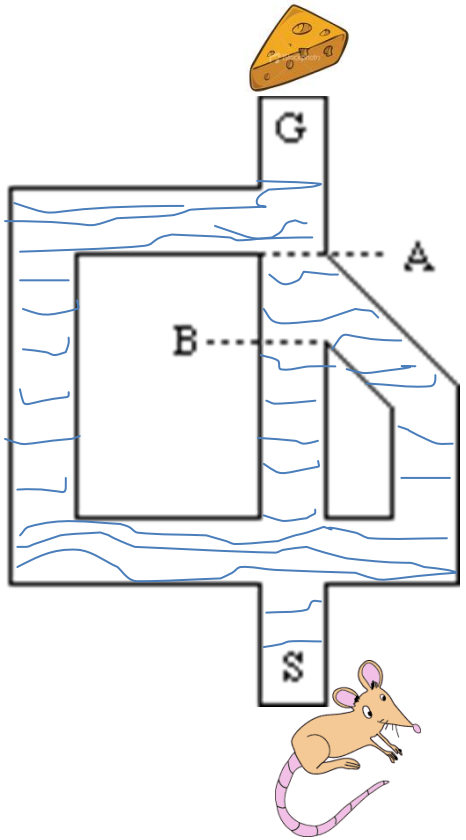



G
A
B
S


G
A
B
S
G
A
B
S


G
A
B
S
G
A
B
S
G
A
B
S

- **Flexibility of behaviour** i.e., behaviour guided by reasoning to the solution of a novel problem
- **Example** Rats tend to take the shortest route through a maze rather than the one that has been most reinforced **(Hull and Tolman)**



AND THE RAT CAN NAVIGATE AFTER A MAZE IS FLOODED – S-R LINKS ARE RADICALLY DIFFERENT

SUGGESTS THE NEED TO PEER INSIDE THE BLACK BOX

- Rats with maps
- Abstract goals
- Human Language
- Beliefs about causality
- Other minds

COMPUTATION AS A FRAMEWORK

COGNITIVE SCIENCE

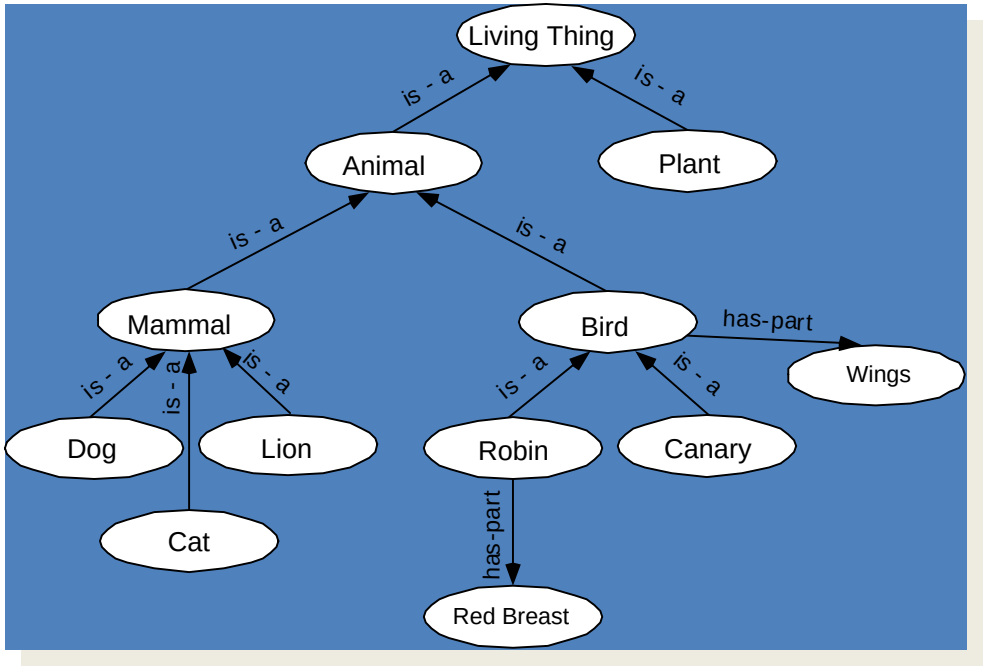
- **Cognitive Science** gradually emerged through the late 40s, 50s, and 60s in the work of Kenneth Craik, George Miller, Jerome Bruner, Herbert Simon, Alan Newell, Noam Chomsky...

- Cognition as computation

Allowed the mind back in to mediate between S and R, in causing intelligent behaviour

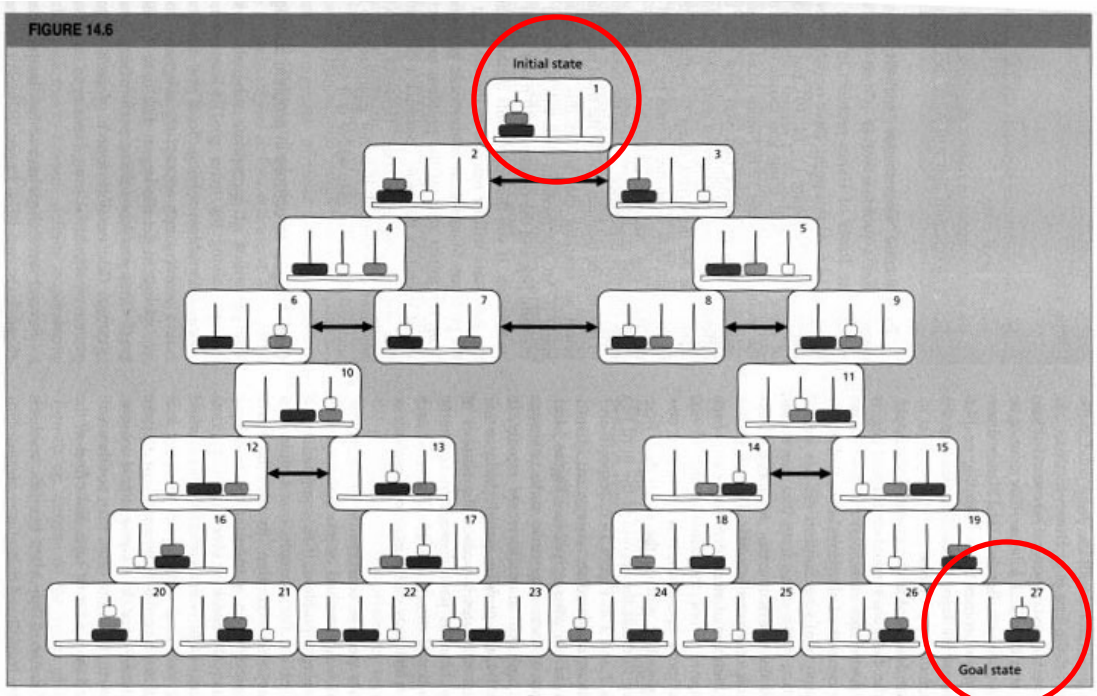
SYMBOLIC KNOWLEDGE REPRESENTATION

- Knowledge organized in semantic networks (Collins & Quillian, 1969)
- Or other logic-based representations, frames, scripts, situation calculus, etc.



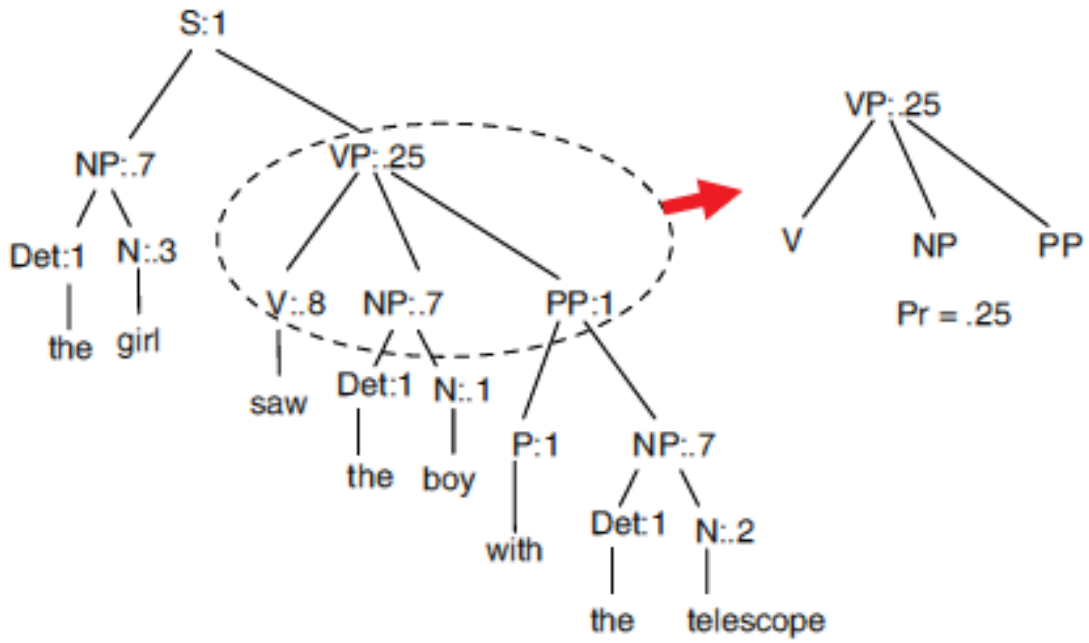
PROBLEM SOLVING AS SYMBOLIC SEARCH

- Problem solving as search in a symbolic problem space (Newell & Simon)



The problem space of legal moves in the Tower of Hanoi problem. If boxes touch each other, or are joined by arrows, this indicates that one can move from one state to the other using a legal operator.

LANGUAGE PROCESSING AS SYMBOL MANIPULATION



(The girl) (saw) (the boy with the telescope)

vs

(The girl) (saw) (the boy) (with the

Two challenges for the symbolic approach

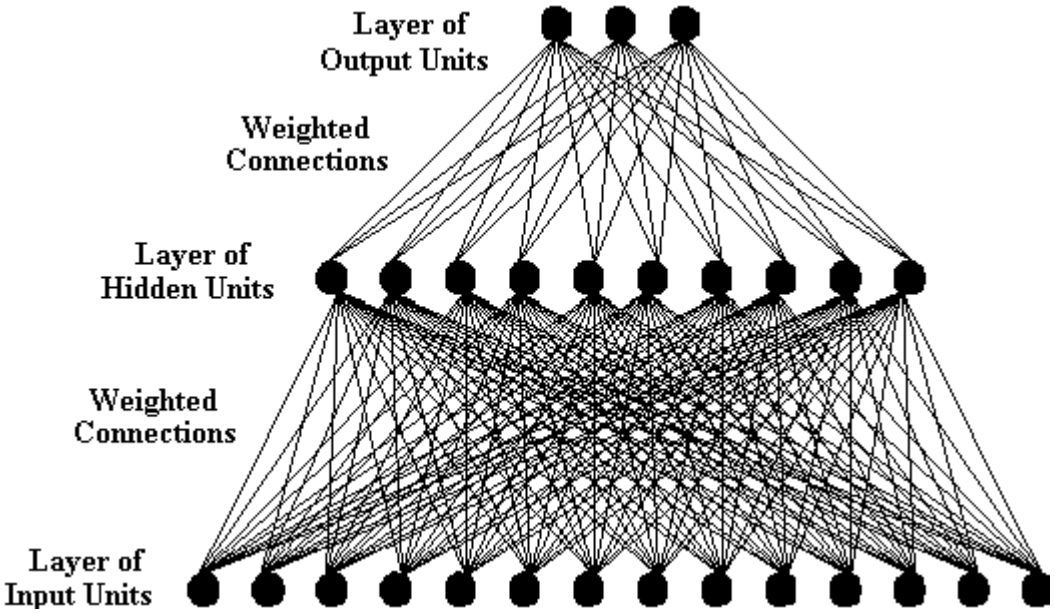
Learning

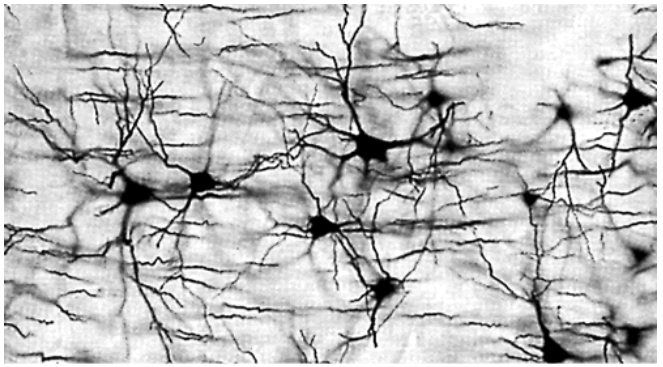
- Knowledge typically hand-coded
- Grammar specified by the linguist
- Inference via logic (or equivalent)—but does not go beyond the information given

Neural plausibility

- 100 step constraint
- Parallel computation
- Numerical values
- How to implement symbolic representations?

CONNECTIONISM





LEARNING VIA ADJUSTING WEIGHTS BY GRADIENT DESCENT (OR SIMILAR);
BUT NO SYMBOLIC REPRESENTATIONS

Which is the right challenge

- To learn symbolic representations, using neurally plausible machinery?

or

- To build a model of cognition without rich symbolic representations (grammars, logic, trees)

3. LEVELS OF EXPLANATION

MARR'S (1982) LEVELS

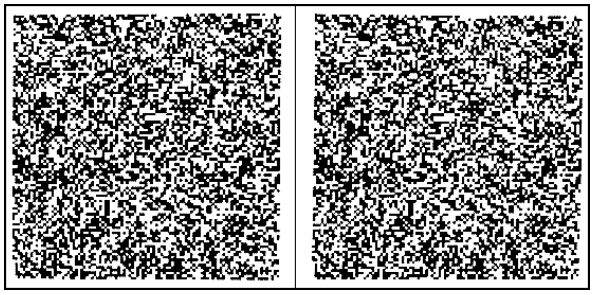
- Computational
  - What problem is the brain solving? What information is required? What is the structure of the environment?
- Algorithmic
  - What algorithms are computed?
    - In general, *not* a direct implementation of calculations the theorist employs at the computational level
- Implementational
  - How are those algorithms implemented

“DISCOVERING” THE POCKET CALCULATOR

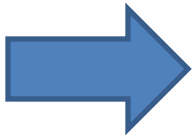
- Computational
- Arithmetic
- Algorithmic
- Decision and binary numbers; algorithms
- Implementational
- Silicon circuits

MATCHING RANDOM DOT STEREOGRAMS

- Computational
- Algorithmic
- Implementational



- The relation between images in each eye: geometry, optics
- Local dot matching algorithms (Mayhew and Frisby)
- Neural circuits



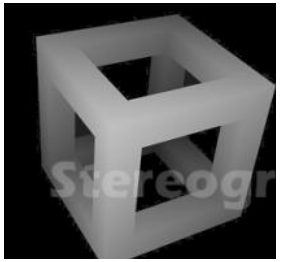


ILLUSTRATION: BAYESIAN COGNITIVE SCIENCE ACROSS THE LEVELS?

- Computational
  - Bayesian picture of structure of reasoning
    - Consistancy
    - Bayesian updating
    - Specific priors

- Algorithmic
  - Graphical models, MCMC learning, etc.

- Implementational
  - Bayesian neural calculations (e.g., Latham, Pouget, Shadlen etc)

SUMMARY

1. Cognitive science meets machine learning

   Flow in information in both directions

2. Historical background: introspection, behaviourism, the computer metaphor

   How does machine learning technology fit in?

3. Levels of explanation

   Which aspects of machine learning carry across at which level?