Thinking Machines

Based on: Chapter 2 of Steven Pinker’s How the Mind Works
Steven Pinker

Cognitive scientist and Harvard Professor

- Thinking Machines
Steven Pinker

- Cognitive scientist and Harvard Professor
- Science popularization

Books:

*Thinking Machines*
Steven Pinker

- Cognitive scientist and Harvard Professor
- Science popularization
Steven Pinker

Cognitive scientist and Harvard Professor
Science popularization
What is the difference between mind and brain?
What is the difference between mind and brain?

1. What is intelligence?
Main Themes

What is the difference between mind and brain?

1. What is intelligence?

2. **Mind** or Software: Computational Theory of the Mind (Abstract rules and representations)
Main Themes

What is the difference between mind and brain?

1. What is intelligence?

2. **Mind** or Software: Computational Theory of the Mind (Abstract rules and representations)

3. **Brain** or Hardware: Neural Networks (neuro-logical networks)
Main Themes

What is the difference between mind and brain?

1. What is intelligence?
2. **Mind** or Software: Computational Theory of the Mind (Abstract rules and representations)
3. **Brain** or Hardware: Neural Networks (neuro-logical networks)
4. Connectionism
What is the difference between mind and brain?

1. What is intelligence?
2. **Mind** or Software: Computational Theory of the Mind (Abstract rules and representations)
3. **Brain** or Hardware: Neural Networks (neuro-logical networks)
4. Connectionism

Till page 73 of pdf file (i.e. page 131 in the chapter)
1. Understand how humans learn and process language
HUL381: Course Objectives

1. Understand how humans learn and process language
2. Examine how computers learn and process language
HUL381: Course Objectives

1. Understand how humans learn and process language
2. Examine how computers learn and process language
3. Use computers to understand how humans process language
Two Fundamental Questions about the Mind

1. What makes intelligence possible?
Two Fundamental Questions about the Mind

1. What makes intelligence possible?
2. What makes consciousness possible?
What is intelligence?
What is intelligence?

1. “Whatever IQ tests measure” (Aptitude)
What is intelligence?

1. “Whatever IQ tests measure” (Aptitude)
2. Rational, human-like thought
What is rationality?
What is rationality?
To make decisions by some set of rules (i.e. grounds of truth)
What is rationality?
To make decisions by some set of rules (i.e. grounds of truth)

1. Correspondance to reality

OR
What is rationality?
To make decisions by some set of rules (i.e. grounds of truth)

1. Correspondance to reality

OR

2. Soundness of inference
Soundness of Inference

Figure: Chess moves
Ability to attain goals by means of decisions based on rational (truth-obeying) rules (Newell and Simon)
Ability to attain goals by means of decisions based on rational (truth-obeying) rules (Newell and Simon)

1. Specifying a goal
Ability to attain goals by means of decisions based on rational (truth-obeying) rules (Newell and Simon)

1. Specifying a goal
2. Assessing the current situation to see discrepancy with the goal
Ability to attain goals by means of decisions based on rational (truth-obeying) rules (Newell and Simon)

1. Specifying a goal
2. Assessing the current situation to see discrepancy with the goal
3. Applying a set of operations to reduce the difference
Ability to attain goals by means of decisions based on rational (truth-obeying) rules (Newell and Simon)

1. Specifying a goal
2. Assessing the current situation to see discrepancy with the goal
3. Applying a set of operations to reduce the difference

Desires pursued using beliefs which are **approximately** or **probabilistically** true
Flesh suffused with non-material soul (theology)
Source of Intelligence

- Flesh suffused with non-material soul (theology)
- Mind comes from extraordinary form of matter
Source of Intelligence

- Flesh suffused with non-material soul (theology)
- Mind comes from extraordinary form of matter
  “brain secretes the mind” (Darwin)
Source of Intelligence

- Flesh suffused with non-material soul (theology)
- Mind comes from extraordinary form of matter “brain secretes the mind” (Darwin)
- Energy flow or force field (Freud)
Source of Intelligence

- Flesh suffused with non-material soul (theology)
- Mind comes from extraordinary form of matter “brain secretes the mind” (Darwin)
- Energy flow or force field (Freud)
- Intelligence arises from *information*! (Turing)
What is information?
What is information?
Correlation between two things produced by a lawful process
What is information?
Correlation between two things produced by a lawful process
(As opposed to chance)
- Tree rings indicate age
What is information?
Correlation between two things produced by a lawful process
(As opposed to chance)

- Tree rings indicate age
- Symbol: Piece of matter carrying information about a certain state of affairs
Information

What is information?
Correlation between two things produced by a lawful process
(As opposed to chance)

▷ Tree rings indicate age
▷ Symbol: Piece of matter carrying information about a certain state of affairs
▷ A machine to scan tree rings and make marks
What guarantee do we have for:

- A collection of symbols for their effects to make sense?
- Effects correspond to some meaningful state of the world?

Guarantee comes from the work of Alan Turing!
Alan Turing

Figure: 1912-1954

- Computer scientist, mathematician, logician, cryptanalyst, philosopher, mathematical biologist, and marathon and ultra distance runner.

Turing was prosecuted in 1952 for homosexual acts.

Died of cyanide poisoning.

His work shortened the WW-II by as many as two to four years.
Computer scientist, mathematician, logician, cryptanalyst, philosopher, mathematical biologist, and marathon and ultra distance runner.

His work shortened the WW-II by as many as two to four years.
Alan Turing

Figure: 1912-1954

- Computer scientist, mathematician, logician, cryptanalyst, philosopher, mathematical biologist, and marathon and ultra distance runner.
- His work shortened the WW-II by as many as two to four years
- Turing was prosecuted in 1952 for homosexual acts

Apple
Alan Turing

Figure: 1912-1954

- Computer scientist, mathematician, logician, cryptanalyst, philosopher, mathematical biologist, and marathon and ultra distance runner.
- His work shortened the WW-II by as many as two to four years
- Turing was prosecuted in 1952 for homosexual acts
- Died of cyanide poisoning
Turing Machine

Universal Turing Machine

State Transition Diagram

Turing Machine Description

Infinite Tape

0 1 0 0 1 1 0 0 0
If the world obeys math equations (solvable step-by-step)
If the world obeys math equations (solvable step-by-step)

A machine to simulate the world and make predictions about it can be built
If the world obeys math equations (solvable step-by-step)
A machine to simulate the world and make predictions about it can be built
To the extent that rational thought corresponds to logic rules
If the world obeys math equations (solvable step-by-step)
A machine to simulate the world and make predictions about it can be built
To the extent that rational thought corresponds to logic rules
A machine to carry out rational thought can be built
If the world obeys math equations (solvable step-by-step)
A machine to simulate the world and make predictions about it can be built
To the extent that rational thought corresponds to logic rules
A machine to carry out rational thought can be built
To the extent that language can be captured by grammatical rules
Rational Machines

- If the world obeys math equations (solvable step-by-step)
- A machine to simulate the world and make predictions about it can be built
- To the extent that rational thought corresponds to logic rules
- A machine to carry out rational thought can be built
- To the extent that language can be captured by grammatical rules
- A machine to produce grammatical sentences can be built
Rational Machines

- If the world obeys math equations (solvable step-by-step)
- A machine to simulate the world and make predictions about it can be built
- To the extent that rational thought corresponds to logic rules
  - A machine to carry out rational thought can be built
- To the extent that language can be captured by grammatical rules
  - A machine to produce grammatical sentences can be built
- If thinking consists of applying well-specified rules, a thinking machine can be built
<table>
<thead>
<tr>
<th>Kinship Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lono-Term Memory</strong></td>
</tr>
<tr>
<td>Abel parent-of Me</td>
</tr>
<tr>
<td>Bella parent-of Me</td>
</tr>
<tr>
<td>Claudia sibling-of Me</td>
</tr>
<tr>
<td>Duddie sibling-of Me</td>
</tr>
<tr>
<td>Edgar sibling-of Abel</td>
</tr>
<tr>
<td>Fanny sibling-of Abel</td>
</tr>
<tr>
<td>Gordie sibling-of Bella</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
Desiderata for Theory of Psychology

Should predict:

- Complex representations for difficult task (compared to an easier task)
Desiderata for Theory of Psychology

Should predict:

- Complex representations for difficult task (compared to an easier task)
- Two similar things have more similar symbols
Desiderata for Theory of Psychology

Should predict:

- Complex representations for difficult task (compared to an easier task)
- Two similar things have more similar symbols
- Salient entities have different representations from their neighbours
How are Concepts Organized?

▶ You are learning to read the word *elk* in a new font.
How are Concepts Organized?

- You are learning to read the word *elk* in a new font
- Do you relearn that the word meaning and that it is a noun?
How are Concepts Organized?

- You are learning to read the word *elk* in a new font
- Do you relearn that the word meaning and that it is a noun?
- If you learn that *wapiti* is a synonym for *elk*, do you relearn everything?
How are Concepts Organized?

- You are learning to read the word *elk* in a new font.
- Do you relearn that the word meaning and that it is a noun?
- If you learn that *wapiti* is a synonym for *elk*, do you relearn everything?
- If you had learned the font as black ink on white paper, would you have to relearn it for white ink on red paper?
Figure: Semantic Network
Mental Representations

- **Visual images**: Template in a two-dimensional, picturelike mosaic
Mental Representations

- **Visual images**: Template in a two-dimensional, picturelike mosaic
- **Phonological representation**: A stretch of syllables
Mental Representations

- **Visual images**: Template in a two-dimensional, picturelike mosaic
- **Phonological representation**: A stretch of syllables
- **Grammatical representation**: Nouns, verbs, phrases, phonemes and syllables arranged into hierarchical trees
Mental Representations

- **Visual images**: Template in a two-dimensional, picturelike mosaic
- **Phonological representation**: A stretch of syllables
- **Grammatical representation**: Nouns, verbs, phrases, phonemes and syllables arranged into hierarchical trees
- **Mentalese**: The language of thought in which our conceptual knowledge is couched

Pinker Thinking Machines
Figure: Units of language
News headline when serial killer Ted Bundy wins a stay on his execution
Language of Thought (Mentalese)

News headline when serial killer Ted Bundy wins a stay on his execution

- *Bundy beats date with chair*
Language of Thought (Mentalese)

News headline when serial killer Ted Bundy wins a stay on his execution

- *Bundy beats date with chair*
- Abstract rules to process mental symbols (representations)
1. Natural Computation NOT Artificial Intelligence
1. Natural Computation NOT Artificial Intelligence
2. Processing of symbols
Computational Theory of the Mind

1. Natural Computation NOT Artificial Intelligence
2. Processing of symbols
3. Arrangements of matter that have both *representational* and *causal* properties
1. Natural Computation NOT Artificial Intelligence
2. Processing of symbols
3. Arrangements of matter that have both *representation*al and *causal* properties
4. If interpretation of input symbols is TRUE, then output symbols also TRUE!
Computational Theory of the Mind

1. Natural Computation NOT Artificial Intelligence
2. Processing of symbols
3. Arrangements of matter that have both *representational* and *causal* properties
4. If interpretation of input symbols is TRUE, then output symbols also TRUE!
5. **Intelligence is computation!**
What makes a system smart?

1. What parts of the machine stand for
What makes a system smart?

1. What parts of the machine stand for
2. How patterns of change inside it mirror truth preserving relationships
What makes a system smart?

1. What parts of the machine stand for
2. How patterns of change inside it mirror truth preserving relationships
3. TRUTH: Absolute or probabilistic or fuzzy
John Searle’s article *Minds, Brains, and Programs* (1980)

Imagine you do not know Chinese
Criticism: Chinese Room Argument

John Searle’s article *Minds, Brains, and Programs* (1980)

- Imagine you do not know Chinese
- You have a rulebook connecting Chinese symbols to other Chinese symbols
Criticism: Chinese Room Argument

John Searle’s article *Minds, Brains, and Programs* (1980)

- Imagine you do not know Chinese
- You have a rulebook connecting Chinese symbols to other Chinese symbols
- For each Chinese symbol input rulebook outputs an answer:
  - Incoming: *two plus two*
  - Rulebook says: *four*
Criticism: Chinese Room Argument

John Searle’s article *Minds, Brains, and Programs* (1980)

- Imagine you do not know Chinese
- You have a rulebook connecting Chinese symbols to other Chinese symbols
- For each Chinese symbol input rulebook outputs an answer:
  - Incoming: *two plus two*
  - Rulebook says: *four*
- Originally intended to argue against Strong AI, where computers have minds
Figure: John Searle’s Chinese Room
Hardware: Neurons

Figure: Courtesy: Hannah Devlin
How to simulate logic gates? (McCullough and Pitts 1943)
How to simulate logic gates? (McCullough and Pitts 1943)

*Logical Calculus of the Ideas Immanent in Nervous Activity*
Neuro-logical Properties

How to simulate logic gates? (McCullough and Pitts 1943)

A Logical Calculus of the Ideas Immanent in Nervous Activity

1. AND
2. OR
3. NOT

Evaluate:

\[ \left( X \text{ chews cud} \right) \text{ and } \left( X \text{ has hooves} \right) \text{ or } \left( X \text{ has fins} \right) \text{ and } \left( X \text{ has scales} \right) \]
Neuro-logical Properties

How to simulate logic gates? (McCullough and Pitts 1943)

**A Logical Calculus of the Ideas Immanent in Nervous Activity**

1. **AND**
2. **OR**
3. **NOT**
4. **Evaluate:**
   
   \[(X \text{ chews cud}) \text{ and } (X \text{ has hooves})\] or \[(X \text{ has fins}) \text{ and } (X \text{ has scales})\]
How to simulate logic gates? (McCullough and Pitts 1943)

A Logical Calculus of the Ideas Immanent in Nervous Activity

1. AND
2. OR
3. NOT
4. Evaluate:

\[(X \text{ chews cud}) \text{ and } (X \text{ has hooves}) \] or \[(X \text{ has fins}) \text{ and } (X \text{ has scales})\]

What about XOR?
Neuron Simulation

\[ a_i = g(\text{in}_i) \]

\[ \sum \text{Input Links} \]

\[ \text{Input Function} \quad \text{Activation Function} \quad \text{Output} \]

Pinker Thinking Machines
Figure: Binary digits for Letter B: 01000010
Patterns of Activity over Sets of Units

Figure: Binary digits for Letter B: 01000010
Neural Networks and Cognition

Good at modelling “fuzziness”

1. Category/set membership
Good at modelling “fuzziness”

1. Category/set membership
2. Partial visual/auditory patterns
Neural Networks and Cognition

Good at modelling “fuzziness”

1. Category/set membership
2. Partial visual/auditory patterns
3. Visual illusions
Which of the following is a vegetable?
Which of the following is a vegetable?

1. Rock (0.0)
Which of the following is a vegetable?

1. Rock (0.0)
2. Ketchup (0.1)
Which of the following is a vegetable?

1. Rock (0.0)
2. Ketchup (0.1)
3. Garlic (0.4)
Category Membership

Which of the following is a vegetable?

1. Rock (0.0)
2. Ketchup (0.1)
3. Garlic (0.4)
4. Cabbage (1.0)
Activation of Meaning Units

green leafy crunchy ... magnetic mobile ...
Auto-Associators

Why not let each unit activate/inhibit other units?
Why not let each unit activate/inhibit other units?
Properties of Auto-Associators

1. Reconstructive and content addressable memory
Properties of Auto-Associators

1. Reconstructive and content addressable memory
2. Graceful degradation
Properties of Auto-Associators

1. Reconstructive and content addressable memory
2. Graceful degradation
3. Constraint satisfaction
Properties of Auto-Associators

1. Reconstructive and content addressable memory
2. Graceful degradation
3. Constraint satisfaction
4. Generalizes automatically
Properties of Auto-Associators

1. Reconstructive and content addressable memory
2. Graceful degradation
3. Constraint satisfaction
4. Generalizes automatically
5. Learn from examples, where learning is a change in the weights.
Graceful Degradation: Partial Percepts

MINE
Constraint satisfaction
Global Ambiguity: Visual Illusion

Is the red dot on the near or far corner?

Figure: Necker cube
Is the red dot on the near or far corner?
Global Ambiguity: Visual Illusion

Is the red dot on the near or far corner?

Figure: Necker cube
Learning from Experience

Figure: Binary Linear Classifier
Separability

Separable

Not Separable
Figure: Pattern associator with a learning technique
1. In 1969 (Minsky & Papert) showed that the perceptron could not learn functions which are not linearly separable (XOR)
Lighthill Report

1. In 1969 (Minsky & Papert) showed that the perceptron could not learn functions which are not linearly separable (XOR)
2. Research into neural networks went into decline throughout the 1970’s.
1. In 1969 (Minsky & Papert) showed that the perceptron could not learn functions which are not linearly separable (XOR)
2. Research into neural networks went into decline throughout the 1970’s.
1. In 1969 (Minsky & Papert) showed that the perceptron could not learn functions which are not linearly separable (XOR)

2. Research into neural networks went into decline throughout the 1970’s.


4. Formed the basis for the decision by the British government to end support for AI research
To solve the XOR problem:
Hidden Layers

To solve the XOR problem:

▶ Make network less of a stimulus-response arrangement
To solve the XOR problem:

- Make network less of a stimulus-response arrangement
- *Internal representation* between input and output layers
To solve the XOR problem:

- Make network less of a stimulus-response arrangement
- *Internal representation* between input and output layers
- XOR: \((A \text{ OR } B) \text{ AND NOT } (A \text{ AND } B)\)
Empiricist Roots of Neural Networks

Locke, Hume, Berkeley and Mill

Experience shows us a number of uniform effects, resulting from certain objects. When a new object, endowed with similar sensible qualities, is produced, we expect similar powers and forces, and look for a like effect. From a body of like color and consistence with bread we expect like nourishment and support. (Hume 1748)
Empiricist Roots of Neural Networks

Locke, Hume, Berkeley and Mill

1. Contiguity of ideas

Experience shows us a number of uniform effects, resulting from certain objects. When a new object, endowed with similar sensible qualities, is produced, we expect similar powers and forces, and look for a like effect. From a body of like color and consistence with bread we expect like nourishment and support. (Hume 1748)
Empiricist Roots of Neural Networks

Locke, Hume, Berkeley and Mill

1. Contiguity of ideas
2. Resemblance of ideas
Empiricist Roots of Neural Networks

Locke, Hume, Berkeley and Mill

1. Contiguity of ideas
2. Resemblance of ideas

Experience shows us a number of uniform effects, resulting from certain objects. When a new object, endowed with similar sensible qualities, is produced, we expect similar powers and forces, and look for a like effect. From a body of like color and consistence with bread we expect like nourishment and support. (Hume 1748)
## Rationalism & Empiricism

<table>
<thead>
<tr>
<th>Empiricism</th>
<th>Rationalism</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Knowledge is based on experience and experimentation.</td>
<td>- Knowledge is based on the use of reason or logic.</td>
</tr>
<tr>
<td>- Experimental science is the paradigm of knowledge.</td>
<td>- Mathematics is the paradigm of knowledge.</td>
</tr>
<tr>
<td>- Experience and experiment rarely, if ever, produce certainty.</td>
<td>- Genuine knowledge is certain.</td>
</tr>
<tr>
<td>- Some empiricists believe that mathematics can be certain.</td>
<td>- Relation to experience:</td>
</tr>
<tr>
<td></td>
<td>- Experience does not produce certainty and does not conform to reason.</td>
</tr>
<tr>
<td></td>
<td>- Thus, experience is at best second-class knowledge.</td>
</tr>
</tbody>
</table>

**Figure:** Courtesy: Prof William Blattner
1. Where do the rules and representations in mentalese leave off and the neural networks begin?
Overview

1. Where do the rules and representations in mentalese leave off and the neural networks begin?
2. Highest level (conscious thought): Rules to plod through symbols?
Overview

1. Where do the rules and representations in mentalese leave off and the neural networks begin?
2. Highest level (conscious thought): Rules to plod through symbols?
3. Lower level (everyday thought): Rules implemented in neural networks?
Overview

1. Where do the rules and representations in mentalese leave off and the neural networks begin?
2. Highest level (conscious thought): Rules to plod through symbols?
3. Lower level (everyday thought): Rules implemented in neural networks?
Views on Human Intelligence

What can account for most of human intelligence?
What can account for most of human intelligence?

1. **Connectionism**: Simple networks by themselves (Rumelhart and McClelland)
What can account for most of human intelligence?

1. **Connectionism**: Simple networks by themselves (Rumelhart and McClelland)

2. **Pinker’s view**: *Structuring* of networks into programs for manipulating symbols
1. Mind is one big hidden-layer network
Connectionism

1. Mind is one big hidden-layer network
2. Intelligence emerges when a trainer, the environment, tunes connection weights

Reason 1: Our networks have more hidden layers between stimulus and response
Reason 2: We live in an environment of other humans who serve as network trainers
Connectionism

1. Mind is one big hidden-layer network
2. Intelligence emerges when a trainer, the environment, tunes connection weights
3. Rules and symbols approximation for millions of streams of activation in neural connections

Reason 1: Our networks have more hidden layers between stimulus and response
Reason 2: We live in an environment of other humans who serve as network trainers

Pinker
Thinking Machines
1. Mind is one big hidden-layer network
2. Intelligence emerges when a trainer, the environment, tunes connection weights
3. Rules and symbols approximation for millions of streams of activation in neural connections
4. Humans smarter than rats!
5. 
   Reason 1: Our networks have more hidden layers between stimulus and response
   Reason 2: We live in an environment of other humans who serve as network trainers
Connectionism

1. Mind is one big hidden-layer network
2. Intelligence emerges when a trainer, the environment, tunes connection weights
3. Rules and symbols approximation for millions of streams of activation in neural connections
4. Humans smarter than rats!
5. **Reason 1**: Our networks have more hidden layers between stimulus and response
1. Mind is one big hidden-layer network
2. Intelligence emerges when a trainer, the environment, tunes connection weights
3. Rules and symbols approximation for millions of streams of activation in neural connections
4. Humans smarter than rats!
5. **Reason 1**: Our networks have more hidden layers between stimulus and response
6. **Reason 2**: We live in an environment of other humans who serve as network trainers
Propounded by Skinner in 1950’s (inspired from Pavlov’s animal psychology)
Behaviourism

- Propounded by Skinner in 1950’s (inspired from Pavlov’s animal psychology)
- Language learning is based on stimulus-response
Propounded by Skinner in 1950’s (inspired from Pavlov’s animal psychology)

- Language learning is based on stimulus-response
- Explicit instruction, correction and reward crucial
Pavlov’s Experiment

Before conditioning

Unconditioned stimulus

Salivation
Unconditioned response

Before conditioning

Neutral stimulus

No salivation
No conditioned response

During conditioning

Salivation
Unconditioned response

After conditioning

Conditioned stimulus

Salivation
Conditioned response

Unconditioned response

Response
Most facts of language not acquired by explicit feedback and reward/punishment

Most facts of language not discussed in grammar books
Behaviourism: Counter evidence

- Most facts of language not acquired by explicit feedback and reward/punishment
- Most facts of language not discussed in grammar books
- We know more than we are aware of:
  The umpires talked to the players. They then left.
Most facts of language not acquired by explicit feedback and reward/punishment

Most facts of language not discussed in grammar books

We know more than we are aware of:

*The umpires talked to the players.* They then left.

*The soldiers fired at the crowd.* They fell down and died.
Most facts of language not acquired by explicit feedback and reward/punishment

Most facts of language not discussed in grammar books

We know more than we are aware of:

*The umpires talked to the players. They then left.*

*The soldiers fired at the crowd. They fell down and died.*

Ability to generalize (*wug* → *wugs*)
Connectionism: Five Difficulties

1. Individuality
Connectionism: Five Difficulties

1. Individuality
2. Compositionality
Connectionism: Five Difficulties

1. Individuality
2. Compositionality
3. Quantification or variable binding
Connectionism: Five Difficulties

1. Individuality
2. Compositionality
3. Quantification or variable binding
4. Recursion
1. Individuality
2. Compositionality
3. Quantification or variable binding
4. Recursion
5. Fuzzy vs. Crisp versions of the same category
Compositionality

- The baby ate the slug
Compositionality

- The baby ate the slug
- The slug ate the baby
Compositionality

- *The baby ate the slug*
- *The slug ate the baby*
Compositionality

- *The baby ate the slug*
- *The slug ate the baby*
No set of weights can directly model:

- Baby same-as baby.
- Baby different-from slug.
- Slug different-from baby.
- Slug same-as slug.
No set of weights can directly model:

- Baby same-as baby.
- Baby different-from slug.
- Slug different-from baby.
- Slug same-as slug.
▶ Every 45 seconds someone in the US sustains a head injury
Quantification

- Every 45 seconds someone in the US sustains a head injury
- Hildegard wants to marry a man with big muscles
Every 45 seconds someone in the US sustains a head injury

Hildegard wants to marry a man with big muscles

You may fool all the people some of the time; you can even fool some of the people all the time; but you can’t fool all of the people all the time
Every 45 seconds someone in the US sustains a head injury
Every 45 seconds someone in the US sustains a head injury

1. Every forty-five seconds there exists an X [who gets injured]
Every 45 seconds someone in the US sustains a head injury
1. Every forty-five seconds there exists an X [who gets injured]
2. There exists an X who every forty-five seconds [gets injured]
Catastrophic Forgetting

1. Add-2 network causes problems with Add-1 (Neal Cohen and Michael McCloskey)
1. Add-2 network causes problems with Add-1 (Neal Cohen and Michael McCloskey)
2. A bat broke the window (Alan Kawamoto)
Embedding propositions to create a hierarchical tree structure of propositions inside propositions
Recursion

Embedding propositions to create a hierarchical tree structure of propositions inside propositions

1. *The baby ate the slug*
Recursion

Embedding propositions to create a hierarchical tree structure of propositions inside propositions

1. *The baby ate the slug*

2. *The father saw the baby eat the slug*
Embedding propositions to create a hierarchical tree structure of propositions inside propositions

1. *The baby ate the slug*
2. *The father saw the baby eat the slug*
3. *I wonder whether the father saw the baby eat the slug*
Recursion

Embedding propositions to create a hierarchical tree structure of propositions inside propositions

1. *The baby ate the slug*
2. *The father saw the baby eat the slug*
3. *I wonder whether the father saw the baby eat the slug*
4. *The father knows that I wonder whether he saw the baby eat the slug*
Embedding propositions to create a hierarchical tree structure of propositions inside propositions

1. The baby ate the slug
2. The father saw the baby eat the slug
3. I wonder whether the father saw the baby eat the slug
4. The father knows that I wonder whether he saw the baby eat the slug
5. I can guess that the father knows that I wonder whether he saw the baby eat the slug
Fuzzy vs. Crisp Categories

Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman
Fuzzy vs. Crisp Categories

Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman

- 13 is a better example of an odd number than 23
Fuzzy vs. Crisp Categories

Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman

- 13 is a better example of an odd number than 23
- Mother is a better example of a female than a comedienne
Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman

- 13 is a better example of an odd number than 23
- Mother is a better example of a female than a comedienne
- A number is either even or odd
Fuzzy vs. Crisp Categories

Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman

- 13 is a better example of an odd number than 23
- Mother is a better example of a female than a comedienne
- A number is either even or odd
- A person must be either male or female
Fuzzy vs. Crisp Categories

Tests by Sharon Armstrong, Henry Gleitman and Lila Gleitman

- 13 is a better example of an odd number than 23
- Mother is a better example of a female than a comedienne
- A number is either even or odd
- A person must be either male or female
- No gray areas!
Chimpanzees like onions. What about gorillas?
Chimpanzees like onions. What about gorillas?

>- All ravens are crows.
>- All crows are birds.
>- All birds are animals.
>- All animals need oxygen.
Generalizing with/without Experience

- Chimpanzees like onions. What about gorillas?
- All ravens are crows.
- All crows are birds.
- All birds are animals.
- All animals need oxygen.
- So do all ravens need oxygen?
1. What is intelligence?
Summary

1. What is intelligence?
   Intelligence as computation

2. **Mind** or Software: Computational Theory of the Mind
Summary

1. What is intelligence?
   Intelligence as computation
2. **Mind** or Software: Computational Theory of the Mind
3. **Brain** or Hardware: Neural networks
Summary

1. What is intelligence?
   Intelligence as computation

2. **Mind** or Software: Computational Theory of the Mind

3. **Brain** or Hardware: Neural networks

4. Connectionism: Structured propositions (hidden layers)