Announcements

• Read Chapter 19 for Thursday – we will discuss it.

• HW 10: Due Tuesday Dec. 6

• Final project: Due Friday Dec. 9

• Volunteers for project demos?
Benford’s law: Distribution of leading digits

<table>
<thead>
<tr>
<th>d</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30.1%</td>
</tr>
<tr>
<td>2</td>
<td>17.6%</td>
</tr>
<tr>
<td>3</td>
<td>12.5%</td>
</tr>
<tr>
<td>4</td>
<td>9.7%</td>
</tr>
<tr>
<td>5</td>
<td>7.9%</td>
</tr>
<tr>
<td>6</td>
<td>6.7%</td>
</tr>
<tr>
<td>7</td>
<td>5.8%</td>
</tr>
<tr>
<td>8</td>
<td>5.1%</td>
</tr>
<tr>
<td>9</td>
<td>4.6%</td>
</tr>
</tbody>
</table>

Table 1. One of the columns gives the land area of political states and territories in km². The other column contains faked data, generated with a random number generator.

<table>
<thead>
<tr>
<th>State/Territory</th>
<th>Real or Faked Area (km²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>645,807</td>
</tr>
<tr>
<td>Albania</td>
<td>28,748</td>
</tr>
<tr>
<td>Algeria</td>
<td>2,381,741</td>
</tr>
<tr>
<td>American Samoa</td>
<td>197</td>
</tr>
<tr>
<td>Andorra</td>
<td>464</td>
</tr>
<tr>
<td>Anguilla</td>
<td>96</td>
</tr>
<tr>
<td>Antigua and Barbuda</td>
<td>442</td>
</tr>
<tr>
<td>Argentina</td>
<td>2,777,409</td>
</tr>
<tr>
<td>Armenia</td>
<td>29,743</td>
</tr>
<tr>
<td>Aruba</td>
<td>193</td>
</tr>
<tr>
<td>Australia</td>
<td>7,682,557</td>
</tr>
<tr>
<td>Austria</td>
<td>83,858</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>86,530</td>
</tr>
<tr>
<td>Bahamas</td>
<td>13,962</td>
</tr>
<tr>
<td>Bahrain</td>
<td>694</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>142,615</td>
</tr>
<tr>
<td>Barbados</td>
<td>431</td>
</tr>
<tr>
<td>Belgium</td>
<td>30,518</td>
</tr>
<tr>
<td>Belize</td>
<td>22,965</td>
</tr>
<tr>
<td>Benin</td>
<td>112,620</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Explanation of Benford’s law
Explanation of Benford’s law

$X = r \times 10^n$  \hspace{1em} (scientific notation)

$1 \leq r < 10$

The leading digit of $X$ is 1 when $1 \leq r < 2$

The leading digit of $X$ is 1 when $2 \leq r < 3$

and so on.

$log_{10} X = log_{10}(r \times 10^n)$

$= log_{10}(r) + log_{10}(10^n)$

$= log_{10}(r) + n$

The leading digit of $X$ is 1 when $1 \leq r < 2$, or when

$n = 0$: \hspace{.5em} 0 \leq log_{10} X < 0.301$

$n = 1$: \hspace{.5em} 1 \leq log_{10} X < 1.301$

$n = 2$: \hspace{.5em} 2 \leq log_{10} X < 2.301$

etc.

(30.1\% of the time)

The leading digit of $X$ is 2 when $2 \leq r < 3$, or when

$n = 0$: \hspace{.5em} 0.301 \leq log_{10} X < 0.477$

$n = 1$: \hspace{.5em} 1.301 \leq log_{10} X < 1.477$

$n = 2$: \hspace{.5em} 2.301 \leq log_{10} X < 2.477$

etc.

(17.6\% of the time)

The leading digit of $X$ is 9 when $9 \leq r < 10$ or when

$n = 0$: \hspace{.5em} .954 \leq log_{10} X < 1$

$n = 1$: \hspace{.5em} 1.954 \leq log_{10} X < 2$

$n = 2$: \hspace{.5em} 2.954 \leq log_{10} X < 3$

etc.

(4.6\% of the time)
Four Principles of Information Processing In Living Systems

*(Complexity: A Guided Tour, Chapter 12)*

1. Global information is encoded as statistics and dynamics of patterns over the systems components.

2. Randomness and probabilities are essential.

3. The system carries out a fine-grained parallel search of possibilities.

4. The system exhibits a continual interplay of bottom-up and top-down processes.
Analogy-Making as a Complex Adaptive System

Melanie Mitchell*
Biophysics Group
Los Alamos National Laboratory


1 Introduction

This paper describes a computer program, called Copycat, that models how people make analogies. It might seem odd to include such a topic in a collection of papers mostly on the immune system. However, the immune system is one of many systems in nature in which a very large collection of relatively simple agents, operating with no central control and limited communication among themselves, collectively produce highly complex, coordinated, and adaptive behavior. Other such systems include the brain, colonies of social insects, economies, and ecologies. The general study of how such emergent adaptive behavior comes about has been called the study of “complex adaptive systems”.

The Copycat program is meant to model human cognition, and its major contribution is to show how a central aspect of cognition can be modeled as the kind of decentralized, distributed complex system described above. In doing so it proposes principles that I believe are common to all complex adaptive systems, and that are particularly relevant to the study of immunology.
Consider the following cognitive activities
• Recognition:
• Recognition:
  – A child learns to recognize cats and dogs in books as well as in real life.
• Recognition:
  – A child learns to recognize cats and dogs in books as well as in real life.
– People can recognize letters of the alphabet, e.g., ‘A’, in many different typefaces and handwriting styles.
People can recognize letters of the alphabet, e.g., ‘A’, in many different typefaces and handwriting styles.
– People can recognize styles of music:
– People can recognize styles of music:

• “That sounds like Mozart”
– People can recognize styles of music:

• “That sounds like Mozart”

• “That’s a muzak version of ‘Hey Jude’”
– People can recognize styles of music:

  • “That sounds like Mozart”

  • “That’s a muzak version of ‘Hey Jude’”

– People can recognize abstract situations:
– People can recognize styles of music:

  - “That sounds like Mozart”
  - “That’s a muzak version of ‘Hey Jude’”

– People can recognize abstract situations:

  - A “Cinderella story”
  - “Another Vietnam”
  - “Monica-gate”
  - “Shop-aholic”
• People make scientific analogies:
• People make scientific analogies:
  – “Biological competition is like economic competition” (Darwin)
• People make scientific analogies:
  
  – “Biological competition is like economic competition” (Darwin)
  
  – “The nuclear force is like the electromagnetic force” (Yukawa)
• People make scientific analogies:
  
  – “Biological competition is like economic competition” (Darwin)
  
  – “The nuclear force is like the electromagnetic force” (Yukawa)
  
  – “The computer is like the brain” (von Neumann)
People make scientific analogies:

- “Biological competition is like economic competition” (Darwin)
- “The nuclear force is like the electromagnetic force” (Yukawa)
- “The computer is like the brain” (von Neumann)
- “The brain is like the computer” (Simon, Newell, etc.)
• People make unconscious analogies
• People make unconscious analogies
  Man: “I’m going shopping for a valentine for my wife.”
• People make unconscious analogies
  Man: “I’m going shopping for a valentine for my wife.”

  Female colleague: “I did that yesterday.”
Main idea is that we understand abstract concepts by analogy with the concrete physical world.
Some examples

• **Happy/Good = up**
  – “My mood has risen since yesterday.”
  – “Things have been going up lately”
  – “I like to get high.”

• **Sad/Bad = down**
  – “My work is going downhill.”
  – “She’s feeling depressed.”
  – “That’s a real downer.”
• **Understanding/Knowledge = seeing / light**
  – “I see what you mean.”
  – “Thanks for enlightening me.”
  – “Can you shed any light on this situation?.”

• **Confusion/Ignorance = blindness/dark**
  – “I can’t see my way through his argument.”
  – “She left me in the dark.”
  – “I have brainfog in the morning.”
• **Love/Attraction = physical force**
  – “I’m falling in love.”
  – “There’s a strong chemistry between us.”
  – “I could feel the electricity between them; there were a lot of sparks.”
  – “He has a lot of magnetism.”
  – “When I’m with him, the atmosphere is charged.”

• **Ideas = food**
  – “I need some time to digest that idea.”
  – “Your idea is half-baked.”
  – “That’s a theory you can sink your teeth into.”
  – “That gave me some food for thought.”
• **Significant/Important** = *large*
  – “He’s a big name in the field.”
  – “He’s a giant among writers.”
  – “What are the big ideas in your field?”
  – “Her accomplishments tower over those of others in her area.”

• **Vitality / Energy** = *a substance ; people = containers*
  – “He’s brimming with spirit.”
  – “She’s overflowing with vitality.”
  – “He’s devoid of energy.”
  – “I’m always drained at the end of the day.”
• Emotional effect = **physical contact**
  – “Getting fired hit him hard.”
  – “That movie bowled me over.”
  – “She made a big impression on me”
  – “That blew me away.”
  – “I was really struck by his sincerity.”

• **Vitality / Energy = a substance ; people = containers**
  – “He’s brimming with spirit.”
  – “She’s overflowing with vitality.”
  – “He’s devoid of energy.”
  – “I’m always drained at the end of the day.”
Copycat: A computer program that models human analogy-making

(Douglas Hofstadter, Melanie Mitchell, Jim Marshall)
Idealizing analogy-making

abc ---> abd

ijk --->
Idealizing analogy-making

abc  --->  abd

ijk  --->  ijl  (replace rightmost letter by successor)
Idealizing analogy-making

abc   --->   abd

ijk   --->   ijl  (replace rightmost letter by successor)

ijd  (replace rightmost letter by ‘d’)

Idealizing analogy-making

abc ---→ abd

ijk ---→ ijl (replace rightmost letter by successor)

ijd (replace rightmost letter by ‘d’)

ijk (replace all ‘c’s by ‘d’s)
Idealizing analogy-making

\[
\begin{align*}
\text{abc} & \quad \rightarrow \quad \text{abd} \\
\text{ijk} & \quad \rightarrow \quad \text{ijl} \\
 & \quad \quad \quad \quad \quad \text{(replace rightmost letter by successor)} \\
\text{ijd} & \quad \rightarrow \quad \text{ijd} \\
 & \quad \quad \quad \quad \quad \text{(replace rightmost letter by ‘d’)} \\
\text{ijk} & \quad \rightarrow \quad \text{ik} \\
 & \quad \quad \quad \quad \quad \text{(replace all ‘c’s by ‘d’s)} \\
\text{abd} & \quad \rightarrow \quad \text{abd} \\
 & \quad \quad \quad \quad \quad \text{(replace any string by ‘abd’)}
\end{align*}
\]
Idealizing analogy-making

abc ---> abd

iijjkk ---> ?
Idealizing analogy-making

abc ---> abd

iijjkk ---> iijjkl

Replace rightmost letter by successor
Idealizing analogy-making

abc ---> abd
ijjkk ---> ?
Idealizing analogy-making

\[
\begin{align*}
\text{abc} & \quad \longrightarrow \quad \text{abd} \\
\text{iijijjkk} & \quad \longrightarrow \quad \text{iijjll}
\end{align*}
\]

Replace rightmost “letter” by successor
Idealizing analogy-making

abc ---> abd
kji ---> ?
Idealizing analogy-making

abc ---> abd

kji ---> kjj

Replace rightmost letter by successor
Idealizing analogy-making

abc  --->  abd

kji  --->  ?
Idealizing analogy-making

abc ---> abd

kji ---> lji

Replace “rightmost” letter by successor
Idealizing analogy-making

abc ---> abd
kji ---> ?
Idealizing analogy-making

abc ---> abd

kji ---> ?
Idealizing analogy-making

abc ---> abd

Replace rightmost letter by “successor”
Idealizing analogy-making

abc \rightarrow abd
mrrjjj \rightarrow ?
Idealizing analogy-making

abc ---> abd
mrrjjj ---> mrrjjk

Replace rightmost letter by successor
Idealizing analogy-making

abc \rightarrow abd
mrrjjj \rightarrow ?
Idealizing analogy-making

abc ---> abd
mrrjjjj ---> ?
1 2 3
Idealizing analogy-making

abc ---> abd
mrrjjj ---> ?
1 2 3
1 2 4
Idealizing analogy-making

abc  --->  abd
mrrjjj  --->  mrrjjjj
1 2  3  1 2  4
Replace rightmost “letter” by successor
Idealizing analogy-making

$abc \longrightarrow abd$

$xyz \longrightarrow ?$
Idealizing analogy-making

abc ---> abd
xyz ---> xya

Replace rightmost letter by successor
Idealizing analogy-making

abc ---> abd

xyz ---> \textcolor{red}{xya} (not allowed)

Replace rightmost letter by successor
Idealizing analogy-making

abc ---> abd
xyz ---> ?
Idealizing analogy-making

abc ---> abd
xyz ---> ?

last letter in alphabet
Idealizing analogy-making

first letter in alphabet

abc ---> abd

last letter in alphabet

xyz ---> ?
Idealizing analogy-making

first letter in alphabet

abc  --->  abd

last letter in alphabet

xyz  --->  ?
Idealizing analogy-making

Replace “rightmost” letter by “successor”
Idealizing analogy-making

*first letter in alphabet*

\[\text{abc} \quad \longrightarrow \quad \text{ab}d\]

\[\text{xyz} \quad \longrightarrow \quad \text{wy}z\]
The Copycat program
(Hofstadter and Mitchell)

- Inspired by collective behavior in complex systems (e.g., ant colonies)
- Understanding and perception of similarity is built up collectively by many independent simple “agents” working in parallel
- Each agent’s job is to explore a possible way of describing an object or relationship between objects
- Each agent has very limited perceptual and communication abilities
- Each agent makes its decisions (about what to explore probabilistically, based on what it perceives in its environment and on its interaction with other agents.
- In this way the resources (agent time) allocated to a possible exploration depends on its promise, as assessed dynamically as exploration proceeds.
- The agents working together produce an “emergent” understanding of the analogy.
Architecture of Copycat
Architecture of Copycat

Concept network
Architecture of Copycat

Concept network

Workspace

a b c ---> a b d
i i j j k k --＞ ?
Architecture of Copycat

Concept network

Workspace

Perceptual agents (codelets)
Architecture of Copycat

Concept network

Workspace

Perceptual agents (codelets)

Temperature

Perceptual agents (codelets)

Concept network

Workspace

Perceptual agents (codelets)

Temperature

Perceptual agents (codelets)
Workspace
a b c --> a b d

m r r j j j j --> ?
Codelet actions

\[ \begin{array}{ccc} 
  a & b & c & \rightarrow & a & b & d \\
  m & r & r & j & j & j & j & \rightarrow & ? \\
\end{array} \]
\[
\begin{align*}
 a & \quad b & \quad c & \quad \rightarrow & \quad a & \quad b & \quad d \\
 m & \quad r & \quad r & \quad j & \quad j & \quad j & \quad j & \quad \rightarrow & \quad ?
\end{align*}
\]
successorship

\[
\begin{align*}
\text{a} & \quad \text{b} & \quad \text{c} & \quad \text{-->} & \quad \text{a} & \quad \text{b} & \quad \text{d} \\
\text{m} & \quad \text{r} & \quad \text{r} & \quad \text{j} & \quad \text{j} & \quad \text{j} & \quad \text{j} & \quad \text{j} & \quad \text{-->} & \quad \? 
\end{align*}
\]
successorship

a b c --> a b d

m r r j j j j --> ?
• Codelets make probabilistic decisions:
• Codelets make probabilistic decisions:
  – What to look at next
• Codelets make probabilistic decisions:
  – What to look at next
  – Whether to build a structure there
• Codelets make probabilistic decisions:
  – What to look at next
  – Whether to build a structure there
  – How fast to build it
• Codelets make probabilistic decisions:
  – What to look at next
  – Whether to build a structure there
  – How fast to build it
  – Whether to destroy an existing structure there
\[
\begin{align*}
\begin{array}{cccc}
\text{a} & \text{b} & \text{c} & \rightarrow \\
\text{m} & \text{r} & \text{r} & \text{j} & \text{j} & \text{j} & \text{j} & \text{j} & \rightarrow \\
\end{array}
\end{align*}
\]
rightmost --> rightmost??
letter --> letter??
a b c --> a b d

m r r j j j j --> ?

rightmost --> rightmost?
letter --> letter?
leftmost --> leftmost??
letter --> letter??
rightmost --> rightmost?
letter --> letter?
leftmost --> leftmost??   rightmost --> rightmost
letter --> letter??       letter --> letter
leftmost --> leftmost??
letter --> letter??

rightmost --> rightmost
letter --> letter
leftmost --> leftmost??
letter --> letter??

rightmost --> rightmost
letter --> letter

rightmost --> rightmost??
letter --> group??
leftmost --> leftmost?
letter --> letter?
rightmost --> rightmost
letter --> group
leftmost --> leftmost?
letter --> letter?
leftmost --> rightmost??
letter --> letter??
rightmost --> rightmost
letter --> group
leftmost --> leftmost?
letter --> letter?
leftmost --> rightmost??
letter --> letter??
rightmost --> rightmost
letter --> group
Concept Network
Part of Copycat’s Concept Network
Concept Network

- Concepts are activated as instances are noticed in workspace.
a b c --> a b d

x y z --> ?
Part of Copycat’s Concept Network

successor
Concept Network

- Activation of concepts feeds back into “top-down” pressure to notice instances of those concepts in the workspace.
successor

a b c --> a b d

x y z --> ?
successor

\[
\begin{align*}
\text{a} & \quad \text{b} & \quad \text{c} & \quad \text{\textarrow} & \quad \text{a} & \quad \text{b} & \quad \text{d} \\
\text{x} & \quad \text{y} & \quad \text{z} & \quad \text{\textarrow} & \quad \text{?}
\end{align*}
\]
successor

\[
\begin{array}{ccc}
  a & b & c \\
  x & y & z
\end{array}
\quad \rightarrow \quad
\begin{array}{cccc}
  a & b & d & d
\end{array}
\]
Concept Network

- Activated concepts spread activation to neighboring concepts.
Concept Network

- Activation of *link* concepts determines current ease of slippages of that type (e.g., “opposite”).
first a b c --> a b d

leftmost x y z --> }

rightmost
The diagram illustrates the concept of 'first' and 'last' in a sequence. Starting from the leftmost element, 'first' moves to the rightmost element, and then 'last' moves to the opposite position of the original 'first'. The sequence changes as follows:

- From 'a b c', 'first' moves to 'a', 'rightmost' to 'd', and 'last' to 'z'.
- From 'x y z', 'first' moves to 'a', 'rightmost' to 'd', and 'last' to 'z'.

The opposite of the sequence is marked with a bright color, highlighting the transformation from 'first' to 'last'.
first → last

a b c → a b d

x y z → ?

opposite

leftmost

first

last

rightmost
first        rightmost
a  b  c        --->        a  b  d

leftmost       last
x  y  z        --->        ?

first --> last

opposite

leftmost

first

last

rightmost
first → last
rightmost → leftmost

opposite

first

last

rightmost
Temperature
Temperature

- Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
Temperature

• Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  – Little organization $\rightarrow$ high temperature
• Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  – Little organization —> high temperature
  – Lots of organization —> low temperature
High temperature

leftmost --> leftmost?
letter --> letter?
rightmost --> rightmost??
letter --> group??
Medium temperature

leftmost --> leftmost?
letter --> letter?

rightmost --> rightmost
letter --> group
Low temperature
Temperature

- Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  - Little organization —> high temperature
  - Lots of organization —> low temperature
Temperature

- Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  - Little organization —> high temperature
  - Lots of organization —> low temperature
- Temperature feeds back to codelets:
Temperature

• Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  – Little organization —> high temperature
  – Lots of organization —> low temperature

• Temperature feeds back to codelets:
  – High temperature —> low confidence in decisions
    —> decisions are made more randomly
Temperature

- Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  - Little organization $\rightarrow$ high temperature
  - Lots of organization $\rightarrow$ low temperature

- Temperature feeds back to codelets:
  - High temperature $\rightarrow$ low confidence in decisions $\rightarrow$ decisions are made more randomly
  - Low temperature $\rightarrow$ high confidence in decisions $\rightarrow$ decisions are made more deterministically
Temperature

• Measures how well organized the program’s “understanding” is as processing proceeds (a reflection of how good the current worldview is)
  – Little organization —> high temperature
  – Lots of organization —> low temperature

• Temperature feeds back to codelets:
  – High temperature —> low confidence in decisions —> decisions are made more randomly
  – Low temperature —> high confidence in decisions —> decisions are made more deterministically

• **Result:** System gradually goes from random, parallel, bottom-up processing to deterministic, serial, top-down processing
Demo
(Metacat, J. Marshall)
Principles that inspired the Copycat program

1. Global information is encoded as statistics and dynamics of patterns over the systems components.

2. Randomness and probabilities are essential

3. The system carries out a fine-grained parallel search of possibilities.

4. The system exhibits a continual interplay of bottom-up and top-down processes.
Applications of Ideas from Copycat


- **Natural language processing** (Gan, Palmer, and Lua, *Computational Linguistics* 22(4), 1996, pp. 531-553)


- **Image understanding** (Mitchell et al.)