Support for Cognitive Science Experiments
Gathering and modeling data in laboratory and classroom experiments

- Model existing psychological experiments
- Collect human data via sketching
- As visual/spatial processing calibrated, provide automatic data analysis facilities
Two Roles in Experiments

• Cognitive simulation platform
  – Including Evans analogy examples
• Gathering & analyzing human data
  – Exporting ink data
  – Interface simplifications
Two Hypotheses about Human Cognition

• Analogy is a central mechanism of reasoning and learning

• Common sense reasoning primarily relies on within-domain analogical reasoning and generalization, organized around qualitative representations
  – Forbus & Gentner, 1997

• Implications:
  – Symbolic, relational representations essential
  – Matching, not chaining
Structure-Mapping Theory (Gentner, 1983)

- Analogy and similarity involve
  - correspondences between *structured* descriptions
    - Feature vectors are inadequate to model human cognition
    - candidate inferences fill in missing structure in target

- Inference is selective. Not all base knowledge is imported

- Also provides account of similarity, metaphor
- Growing body of psychological evidence that same processes are used in perception, problem solving, and conceptual change
SME: Structure-Mapping Engine

Inputs = propositional descriptions, w/ incremental updates
Output = one or two mappings

Base

Operates in polynomial time, by exploiting graph labels & greedy algorithms

Target

Mappings = correspondences + structural evaluation + candidate inferences
Computational Properties of SME

• Scalable
  – Cases can contain thousands of propositions
  – Cases can be dynamically constructed and expanded from knowledge base contents

• Flexible
  – Has been used with large knowledge bases developed by others (e.g., Cyc, KM)

• Supports Integration
  – Analogy ontology enables smooth integration with logical reasoners

• SME remains the only general-purpose cognitive simulation of analogical matching used with multiple knowledge systems and reasoners
SME can operate over visual structure

Corresponds to what people choose in fast response-time task

Corresponds to what people choose when given more time
Building Blocks for Analogical Processing

- **SME** = Matching
- **MAC/FAC** = Retrieval
- **SEQL** = Generalization w/probabilities
Psychological evidence (examples)

• Used to model existing findings
  – e.g., SME models effects of relational structure on similarity
  – e.g., MAC/FAC models dissociation between surface effects on reminding versus preference for deep structure in mapping
  – e.g., SEQL-based model of Marcus experiment still only one that learns in same span of stimuli as infants, and can handle noise

• Used to predict new findings
  – e.g., SME: Initial stage of metaphor processing is symmetric
  – e.g., SEQL: Can generate orders of presentation which can help/hurt concept learning

• A number of aspects not yet modeled
  – e.g., Working memory capacity limits
Some CogSketch Simulation Examples

Geometric Analogy

- Problems of the form “A is to B as C is to __?”

Raven’s Progressive Matrices

- Used to measure intelligence
- Extensive data on human performance available

Learning spatial prepositions

Best Generalization IN
Size: 3
(candle in bottle, cookie in bowl, marble in water)

- DEFINITE FACTS:
  (rcc8-TPP figure ground)
- POSSIBLE FACTS:
  33% (Basin ground)
  33% (Bowl-Generic ground)
Geometric Analogy

- Problems of the form “A is to B as C is to __?”

A

B

C

1

2

3

4

5
Geometric Analogy

- Problems of the form “A is to B as C is to __?”

A

B

C

1

2

3

4

5
Two-Stage Structure Mapping

First Stage: Differences

A → SME → Δ(A,B)

B → SME → Δ(A,B)

C → SME → Δ(C,1)

1 → SME → Δ(C,1)

2 → SME → Δ(C,2)

3 → SME → Δ(C,3)

4 → SME → Δ(C,4)

5 → SME → Δ(C,5)
Two-Stage Structure Mapping

First Stage: Differences

Second Stage: Similarity Score

Answer 1
Answer 2
Answer 3
Answer 4
Answer 5
Two-Stage Structure Mapping

First Stage: Differences

Second Stage: Similarity Score

Answer 1
Answer 2
Answer 3
Answer 4
Answer 5
Evaluation

• Constructed 20 problems originally used by Evans (1968) with his landmark ANALOGY system
  – Gave problems to 34 participants
  – Ran problems on computational model
• Model chooses answer preferred by humans on all 20 problems
• Model shows a .75 correlation with human timing data
Raven’s Progressive Matrices

- Used to measure intelligence
- Extensive data on human performance available
Two-Stage Structure Mapping

Two-stage mapping process illustrated with two stages of SME (Structural Mapping Elements) and four answers: Answer 1, Answer 2, Answer 3, and Answer 4.
Solving the RPM

- Row solution
- A is to B, as C is to __?
- difference(A,B)
  - In A, the inner object is on the left side of the outer object
  - In B, the inner object is on the right side of the outer object
Solving the RPM

- Column solution
- A is to B, as C is to __?
- difference(A,B)
  - In A, the inner object is on the top half of the outer object
  - In B, the inner object is on the bottom half of the outer object
Evaluation

• Initially evaluated on two fairly easy sections of the standard RPM
  – B: 2x2 matrices, 6 possible answers
  – C: 3x3 matrices, 8 possible answers

• Performed at the level of the average American adult on those sections
  – B: 12/12
  – C: 10/12
Visual Oddity Task
Visual Oddity Task

The image contains a visual oddity task with five rectangles and one diamond. The diamond is the odd one out.
Our model for the Oddity Task

Test: Is one image noticeably less similar to the generalization?
Results

- Overall Performance: 39 / 45
- Correlation w/Americans: 0.656 (aged 8-13)
- Correlation w/Mundurukú: 0.493 (all ages)
- Model can be used to identify operations that contribute to problem difficulty
Copy/Paste from PowerPoint

- Import shapes drawn in PowerPoint via copy/paste
- PowerPoint shape -> Glyph

N.B.: It can take a few moments for the ink processor to finish before the glyphs appear
Copy/Paste from PowerPoint

• What is supported
  – Most simple shapes
    • Straight/Curved, Open/Closed, Custom-drawn, etc
  – Line thickness, line color, fill color
  – Group PowerPoint shapes together to make them a single CogSketch glyph

• Not supported
  – More complex shapes
    • 3D shapes, shapes with multiple polygons, arrows
  – More complex attributes (shading, textures, etc)
Exporting Ink

Comma separated value files can be directly read by modern spreadsheets.

Can pick which bundles are included.
## Result of Ink Export

Note: Timestamp information provided for every ink point in the sketch

<table>
<thead>
<tr>
<th>Sketch Namestring</th>
<th>Sketch Case</th>
<th>Bundle Namestring</th>
<th>Bundle Case</th>
<th>Layer Namestring</th>
<th>Glyph Names Glyph Name string</th>
<th>Ink ID</th>
<th>Ink Type</th>
<th>Ink Point X</th>
<th>Ink Point Y</th>
<th>Point stamp (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shopping Cart</td>
<td>Case-3429195339 Anatomy</td>
<td>&quot;Physical&quot; ObjectL-225 &quot;Handle&quot; Object-154</td>
<td>BCASE-342919545</td>
<td>sketch-polylin</td>
<td>2.2447</td>
<td>0.4583252.35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redux</td>
<td>Case-3429195339 Anatomy</td>
<td>&quot;Physical&quot; ObjectL-225 &quot;Handle&quot; Object-154</td>
<td>BCASE-342919545</td>
<td>sketch-polylin</td>
<td>2.2447</td>
<td>0.4583252.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Case-3429195339 Anatomy</td>
<td>&quot;Physical&quot; ObjectL-225 &quot;Handle&quot; Object-154</td>
<td>BCASE-342919545</td>
<td>sketch-polylin</td>
<td>2.2447</td>
<td>0.4687252.67</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Case-3429195339 Anatomy</td>
<td>&quot;Physical&quot; ObjectL-225 &quot;Handle&quot; Object-154</td>
<td>BCASE-342919545</td>
<td>sketch-polylin</td>
<td>2.2656</td>
<td>0.4791252.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

...
Viewing Timing Data

- Glyph order
  - Need to be in experimenter mode
- Ink replay
Gathering Sketch Data

• Potential CogSketch advantages in data collection:
  – Easier to archive and transmit bits than dead trees
  – Captured digital ink is potentially easier to analyze than video or scanned bitmaps
    • Timing data automatically captured
  – Conceptual labeling could reduce subsequent hand-coding of data ("what’s that??")
  – Visual/spatial processing could become calibrated enough with human judgments to automate some kinds of data analysis

• Potential disadvantages:
  – Teaching participants how to sketch with it
  – Overhead of conceptual labeling can be distracting
Simplifying Concept Labeling

- Worksheets use a simple list
- Can be done in any order
Experimental: Free-form NL input

Allows users to type any string

Currently we cannot analyze such labels
Coming up: Skins and Scripting

• Skins = ability to hide/expose capabilities in the interface
  – Often useful for participants to have fewer distractions
  – You can already choose skins when building worksheets
  – File format and documentation under development to let experimenters generate their own skins.

• Scripting
  – Want ability to run participants through a number of sketching exercises, with minimal or no experimenter intervention
  – Suggestions about what you need would be welcome
Questions?