Reasoning with Visual Depictive Representations

Soar Spatial & Visual Imagery (Soar+SVI)
OUTLINE

• BACKGROUND

• MOTIVATING EXAMPLE

• DEPICTIVE MANIPULATIONS

• RESULTS

• NUGGETS / COAL
WHAT IS MENTAL IMAGERY?

MENTAL IMAGERY

SPATIAL
- Direction
- Distance
- Orientation
- Size
- Topology

VISUAL
- Features
- Color
- Spatial

Specific Shape

Generic Shape
WHY RESEARCH MENTAL IMAGERY?

- **Cognitive Architectures**
  - Amodal, symbolic representations & computations
  - Little *reasoning* with perceptual-based representations (Barsalou 1999, 2008)

- **Resulting in...**
  - Ad-hoc reasoning in tasks rich with spatial and visual properties
  - “Bolted-on” task-dependent components

- **What we want**
  - Link *perceptual-based thought* and *cognition*
  - “Best of both worlds” *multi-representational*, task-*independent* approach
    - Additional functionality
    - Computational advantage
<table>
<thead>
<tr>
<th>Representation</th>
<th>Processing</th>
<th>Uses</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Symbolic</strong></td>
<td>Symbolic manipulation</td>
<td>General Reasoning</td>
<td><strong>object(can)</strong> feature (can, curve) color(can, yellow) on(can, box)</td>
</tr>
<tr>
<td>Amodal</td>
<td></td>
<td>Qualitative Spatial &amp; Visual Reasoning Control imagery operations</td>
<td><strong>object(box)</strong> feature(box, corner) color(box, green) on(can, box)</td>
</tr>
<tr>
<td><strong>Quantitative spatial</strong></td>
<td>Mathematical manipulation</td>
<td>Spatial Reasoning (Generic Shapes)</td>
<td><strong>can</strong> location &lt;2,1,2&gt; orientation 0 height 5 radius 2 box location &lt;0,0,0&gt; orientation 0 length 10 width 6 height 4</td>
</tr>
<tr>
<td>Amodal/Perceptual-based</td>
<td>Mathematical manipulation</td>
<td>Spatial Reasoning (Generic Shapes)</td>
<td><strong>can</strong> location &lt;2,1,2&gt; orientation 0 height 5 radius 2 box location &lt;0,0,0&gt; orientation 0 length 10 width 6 height 4</td>
</tr>
<tr>
<td><strong>Visual depictive</strong></td>
<td>Mathematical manipulation</td>
<td>Spatial Reasoning (Specific Shapes)</td>
<td><strong>Visual Feature Recognition</strong></td>
</tr>
<tr>
<td>Modal/Perceptual-based</td>
<td>Depictive manipulation</td>
<td></td>
<td><strong>Visual Feature Recognition</strong></td>
</tr>
</tbody>
</table>
SCOUT DOMAIN

• Two scouts
  ➢ Leader (Agent)
  ➢ Teammate (Scripted)
  ➢ Goal: maintain visual contact with approaching enemy to “paint picture”

• Task Characteristics
  ➢ Spatial (e.g. relationships between entities, terrain, etc.)
  ➢ Visual (e.g. terrain topological shape)
  ➢ Perceive/Imagine/Re-perceive
  ➢ Planning during execution
  ➢ Merge multiple sources of information (visual, messages, doctrine)

<table>
<thead>
<tr>
<th>Percepts</th>
<th>Actions</th>
<th>Task Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agent’s location and orientation</td>
<td>Look-at scene / map</td>
<td>Enemy doctrine</td>
</tr>
<tr>
<td>Objects</td>
<td>Send message</td>
<td>No-go/go terrain (pixel values)</td>
</tr>
<tr>
<td>Egocentric direction &amp; distance</td>
<td>Turn</td>
<td>Hypothesized key terrain</td>
</tr>
<tr>
<td>Messages from teammate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Actual Situation

Teammate’s View

Agent’s View

Imagined Situation
SCOUT DOMAIN
(Example)

- Agent imagines what it and teammate can see (field of view)

LEGEND:
- Scout
- Enemy
SCOUT DOMAIN

(VIDEO)
DEPICTIVE MANIPULATIONS
(1 of 2)

• **Pixel-level rewrites (Furnas)**
  - Production-like rules (LHS/RHS)
  - Shared image (i.e. working memory)
  - LHS and RHS are pixel patterns
  - Conflict resolution scheme (sequencing)
  - Processing may match other orientations
  - Active manipulation of shapes

• **Contrast with other image processing**
  - Computer vision
    - Sentential manipulations
    - Filters (e.g. Gaussian) where each pixel is rewritten as a specific function of its neighbors’ values
  - Cellular Automata
    - Finite State Machines
    - Next state of a cell based on current state and state of its neighbors
    - Rather than rewriting many local configurations at once

- **“2x2 reduce”**
- **“nibble up”**
- **“nibble left”**
Layer 0-1

- Take specific shape of no-go terrain and obstacles (yellow) into account
- Attention window shift controlled by Soar
- Rules sent from Soar to manipulator via operator elaborations
- Determine approximate path coverage by overlaying each scout’s view frustum
- Simulate alternate view orientations to determine “best” path coverage
DEPICTIVE MANIPULATIONS
(Soar only)

(VIDEO)
DEPICTIVE MANIPULATIONS (Soar+SVI) (VIDEO)
DEPICTIVE MANIPULATIONS
(Soar+SVI)

(VIDEO)
COMPUTATIONAL EFFICIENCY

- Task specific/representation specific
- Task independent/representation specific
- Task Independent/representation independent
\[ I_e = \begin{cases} -1 & \text{if no observation} \\ 1 - \delta & \text{otherwise} \end{cases} \]

\[ \delta = \sqrt{(\text{obs}_x - \text{act}_x)^2 + (\text{obs}_y - \text{act}_y)^2} / \text{d_{acceptable}} \]

\[ \text{d_{acceptable}} = \text{the acceptable square distance} = \sqrt{d_x^2 + d_y^2} \quad \text{where} \quad d_x = d_y = 500 \text{ meters} \]
PROBLEM-SOLVING QUALITY
(Scenario-2)
PROBLEM-SOLVING QUALITY
TOTAL OBSERVATIONS

![Bar chart showing observations for different enemies and conditions.]

- Observations:
  - Enemy-1
  - Enemy-2
  - Enemy-3

- Conditions:
  - Soar+SVI
  - Soar-SVI
  - Observer
• NUGGETS
  ▪ Usefulness of reasoning with visual depictive representation and specialized processing

  ▪ Integration of spatial and visual imagery functional constraints
    o Task-independent cognitive architecture
    o Reasoning emerges from the combination of the representations

• COAL
  ▪ More research required on integration between low-level visual perception and system
  ▪ Unclear how depictive rules are learned