Multi-representational Architectures: Incorporating Visual Imagery

into a

Cognitive Architecture

Soar Visual Imagery (SVI)

Scott Lathrop

John Laird

27th SOAR WORKSHOP

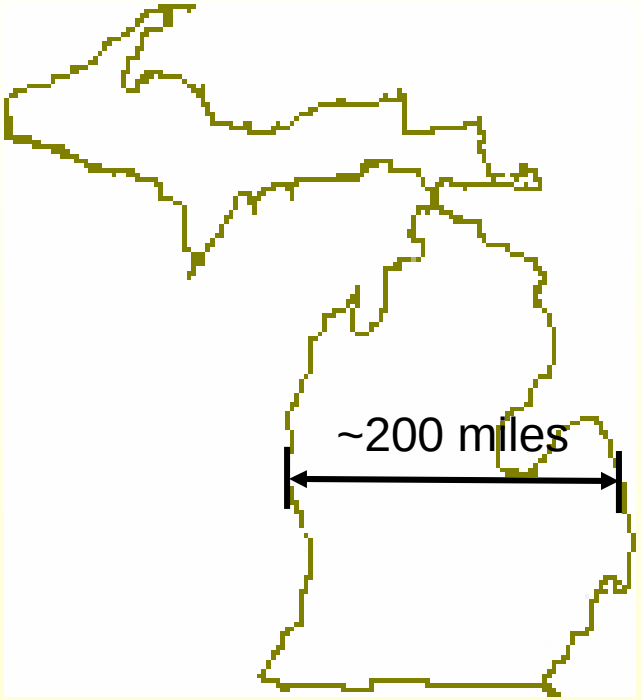
OUTLINE

- REVIEW
- CURRENT ARCHITECTURE (SVI)
- EXPERIMENTAL RESULTS
- FUTURE WORK

WHAT IS VISUAL IMAGERY?

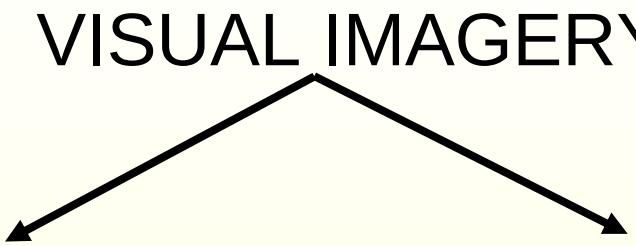
- What city is closer to Ann Arbor: South Bend, Indiana or Columbus, Ohio?

WHAT IS VISUAL IMAGERY



- **What is wider in the center: the lower peninsula of Michigan or Ohio?**

WHAT IS VISUAL IMAGERY?



VISUAL-SPATIAL

- Location, orientation
- Sentential, quantitative representations
- Linear algebra and computational geometry algorithms

VISUAL-DEPICTIVE

- Shape, color, topology, spatial properties
- Depictive, pixel-based representations
- Image algebra algorithms
  - ❑ Sentential/Algebraic algorithms
  - ❑ Depictive/Ordinal algorithms

MULTI-VISUAL REPRESENTATIONS

Representation	Processing	Uses	Example
Abstract symbols	Symbolic manipulation	Qualitative Visual & Spatial Reasoning	object(Mich) object(Ohio) south(Ohio,Mich) in(AA,Mich), sw(AA,Mich) center(Columbus, Ohio) etc.
Hybrid abstract and quantitative symbols	Sentential, algebraic manipulation	Quantitative Spatial Reasoning Intermediate Layer	Michigan shape: rectangle location: <10,20,0> AA shape: point location <20,2,0> Ohio shape: square location: <15,-5,0>
Iconic / Depictive symbols	Algebraic or Depictive manipulation	Visual Feature Recognition Quantitative Spatial Reasoning	

Abstract
Non-commital

Committal
Concrete

REVIEW SUMMARY

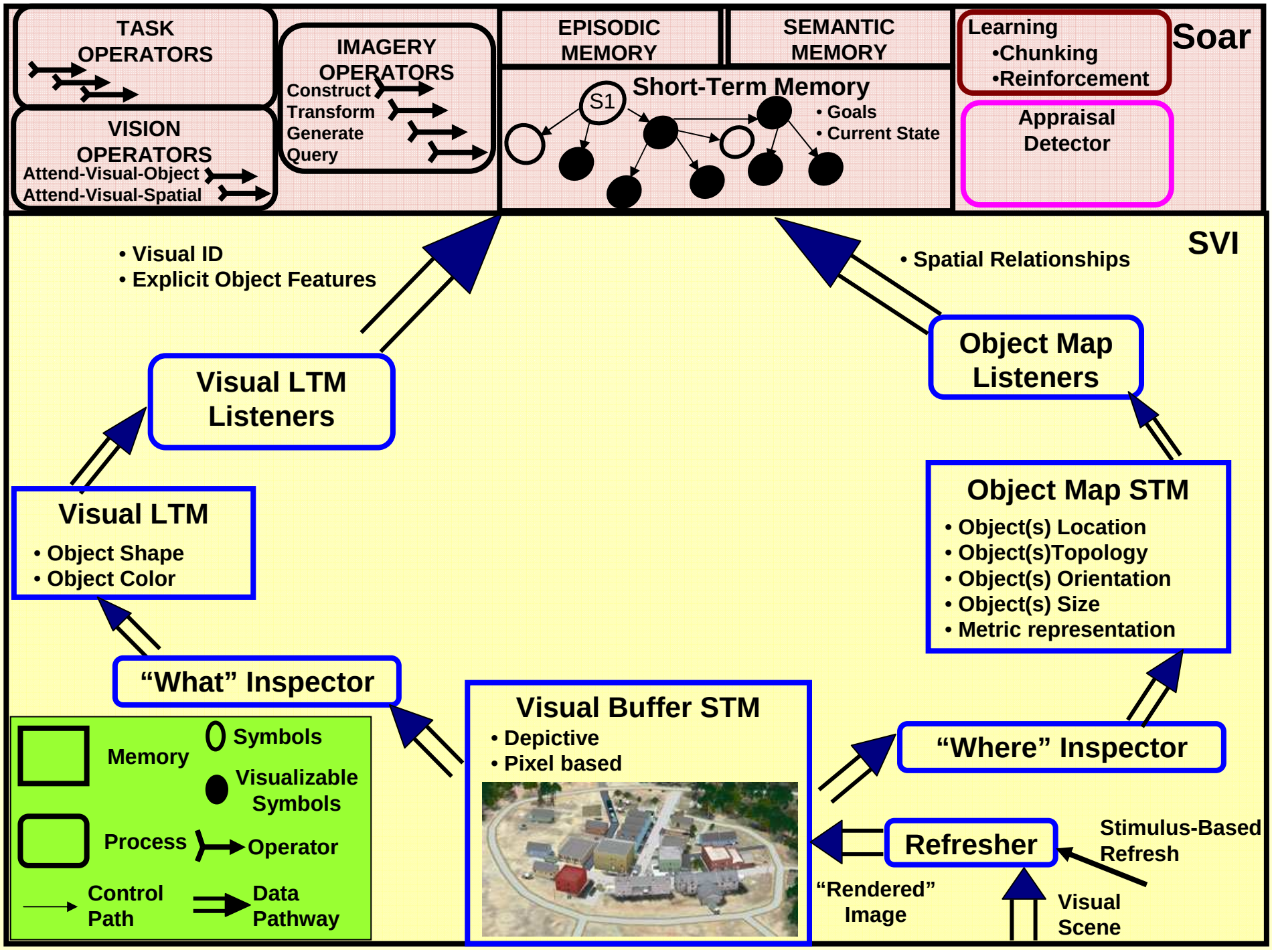
- Why research visual imagery?
  - "Best of both worlds" multi-representational approach
    - Abstract symbolic representations & computations
    - Perceptually-based quantitative and depictive representations
  - Add new capability
    - Visual-spatial reasoning
    - Visual-feature retrieval and reasoning
  - Gain computational advantage

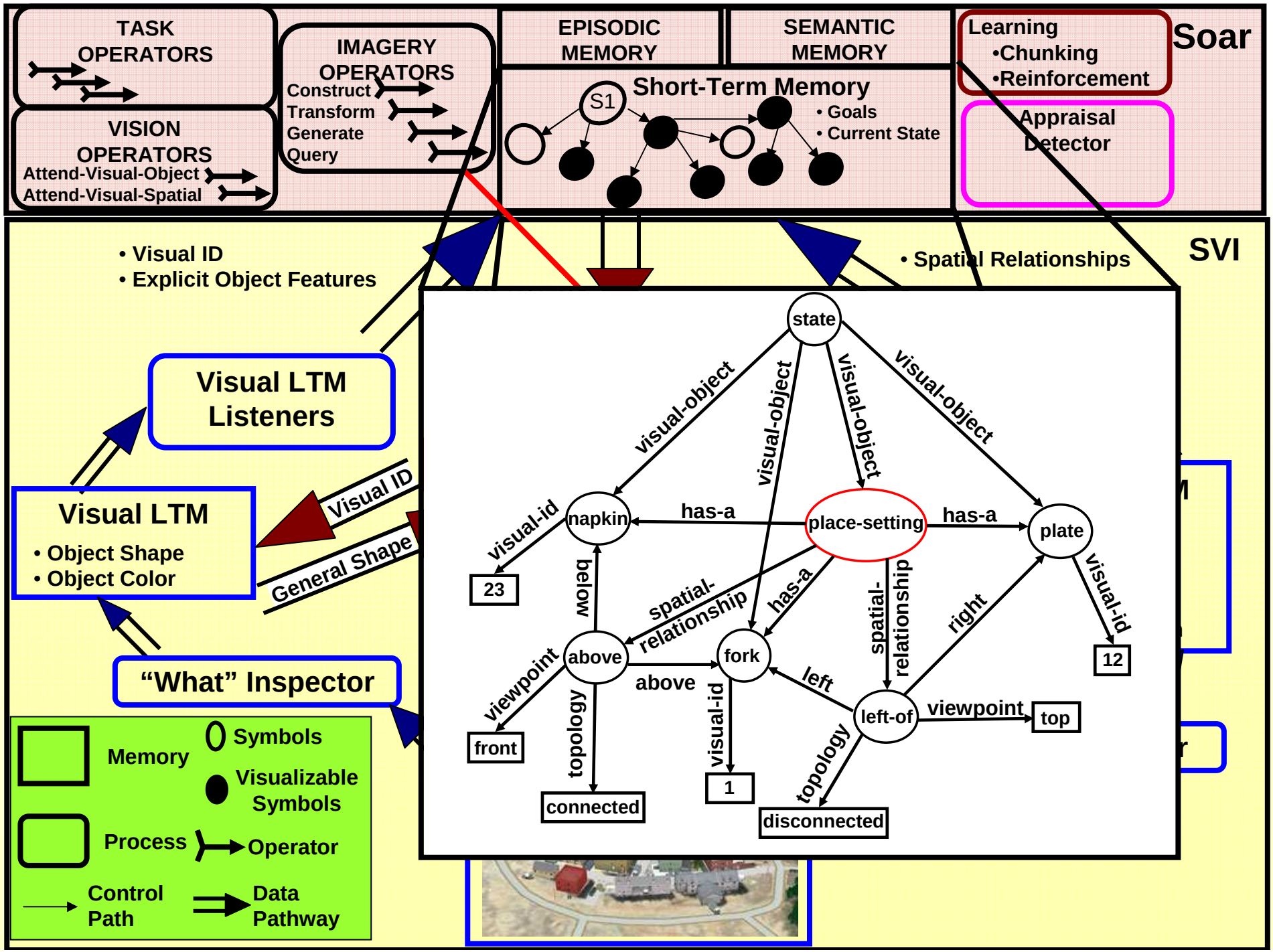
- Previous architecture and experiments focused exclusively on quantitative representations and visual-spatial type tasks

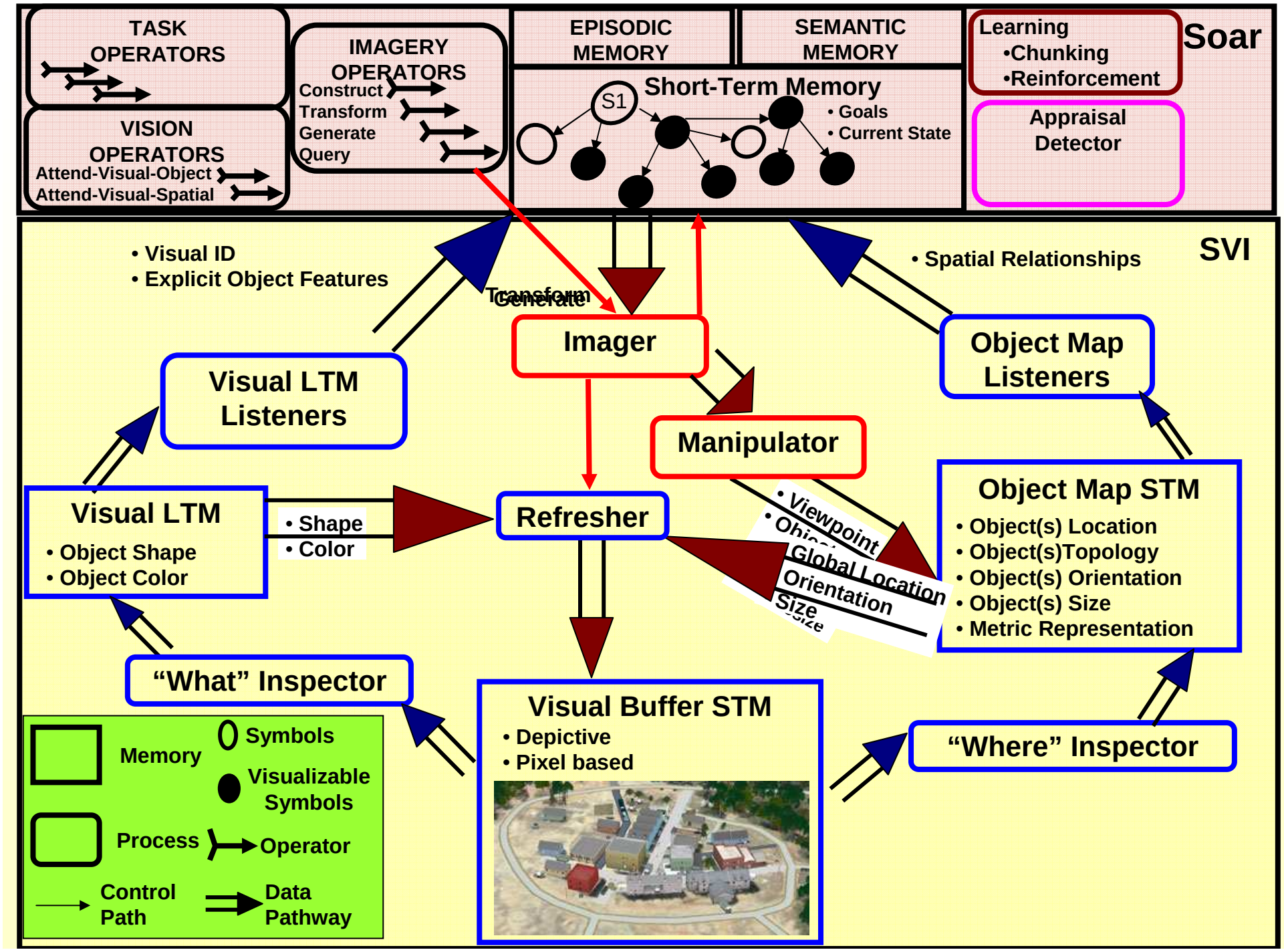
- Open research questions as of last Soar workshop
  - What is a visual image's internal representation? Is there more than one format/data structure?
  - What is the relationship between high-level vision and visual imagery and how does that constrain the architecture?

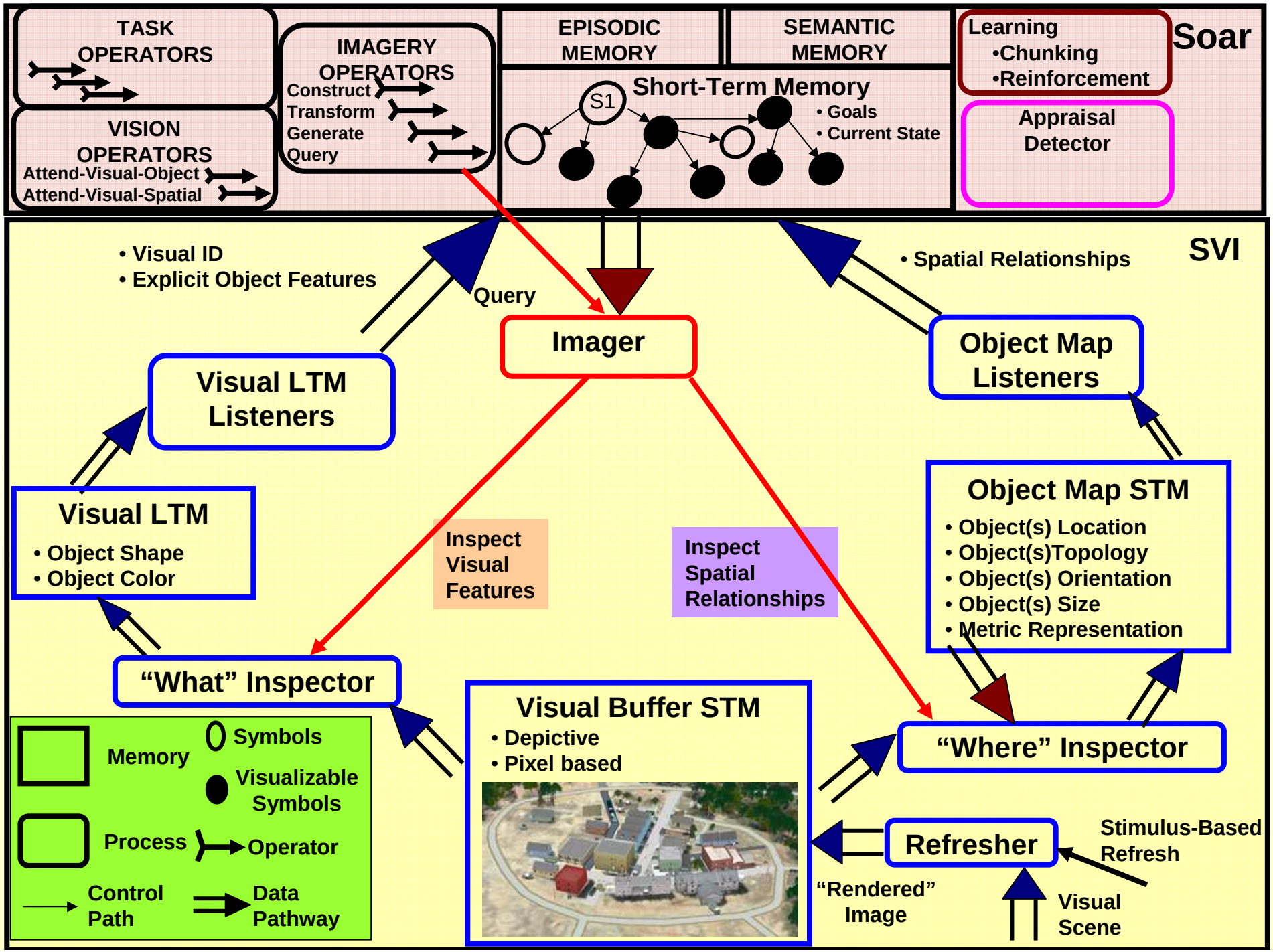
OUTLINE

- REVIEW
- CURRENT ARCHITECTURE (SVI)
- EXPERIMENTAL RESULTS
- FUTURE WORK


TASK OPERATORS
VISION OPERATORS
Attend-Visual-Object
Attend-Visual-Spatial
IMAGERY OPERATORS
Construct
Transform
Generate
Query
EPISODIC MEMORY
SEMANTIC MEMORY
Short-Term Memory
S1
• Goals
• Current State
Learning
•Chunking
•Reinforcement
Appraisal Detector
Soar
SVI
• Visual ID
• Explicit Object Features
• Spatial Relationships
Visual LTM Listeners
Object Map Listeners
Visual LTM
• Object Shape
• Object Color
Object Map STM
• Object(s) Location
• Object(s)Topology
• Object(s) Orientation
• Object(s) Size
• Metric representation
"What" Inspector
Visual Buffer STM
• Depictive
• Pixel based
"Where" Inspector
Refresher
Stimulus-Based Refresh
"Rendered" Image
Visual Scene
Memory
Symbols
Visualizable Symbols
Process
Operator
Control Path
Data Pathway


TASK OPERATORS
IMAGERY OPERATORS
Construct
Transform
Generate
Query
VISION OPERATORS
Attend-Visual-Object
Attend-Visual-Spatial
EPISODIC MEMORY
SEMANTIC MEMORY
Short-Term Memory
S1
• Goals
• Current State
Learning
•Chunking
•Reinforcement
Appraisal Detector
Soar
• Visual ID
• Explicit Object Features
• Spatial Relationships
SVI
Visual LTM Listeners
Visual LTM
• Object Shape
• Object Color
Visual ID
General Shape
"What" Inspector
state
visual-object
napkin
place-setting
plate
has-a
visual-id
23
below
spatial-relationship
above
fork
left-of
viewpoint
front
topology
connected
1
disconnected
top
12
right
Memory
Symbols
Visualizable Symbols
Process
Operator
Control Path
Data Pathway


TASK OPERATORS
VISION OPERATORS
Attend-Visual-Object
Attend-Visual-Spatial
IMAGERY OPERATORS
Construct
Transform
Generate
Query
EPISODIC MEMORY
SEMANTIC MEMORY
Short-Term Memory
S1
• Goals
• Current State
Learning
•Chunking
•Reinforcement
Appraisal Detector
Soar
SVI
• Visual ID
• Explicit Object Features
• Spatial Relationships
Transform
Generate
Imager
Visual LTM Listeners
Object Map Listeners
Manipulator
Visual LTM
• Object Shape
• Object Color
• Shape
• Color
Refresher
• Viewpoint
Global Location
Orientation
Size
Object Map STM
• Object(s) Location
• Object(s)Topology
• Object(s) Orientation
• Object(s) Size
• Metric Representation
"What" Inspector
Visual Buffer STM
• Depictive
• Pixel based
"Where" Inspector
Memory
Symbols
Visualizable Symbols
Process
Operator
Control Path
Data Pathway


TASK OPERATORS
VISION OPERATORS
Attend-Visual-Object
Attend-Visual-Spatial
IMAGERY OPERATORS
Construct
Transform
Generate
Query
EPISODIC MEMORY
SEMANTIC MEMORY
Short-Term Memory
S1
• Goals
• Current State
Learning
•Chunking
•Reinforcement
Appraisal Detector
Soar
SVI
• Visual ID
• Explicit Object Features
Query
• Spatial Relationships
Imager
Visual LTM Listeners
Object Map Listeners
Visual LTM
• Object Shape
• Object Color
Inspect Visual Features
Inspect Spatial Relationships
Object Map STM
• Object(s) Location
• Object(s)Topology
• Object(s) Orientation
• Object(s) Size
• Metric Representation
"What" Inspector
Visual Buffer STM
• Depictive
• Pixel based
"Where" Inspector
Refresher
Stimulus-Based Refresh
"Rendered" Image
Visual Scene
Memory
Symbols
Visualizable Symbols
Process
Operator
Control Path
Data Pathway

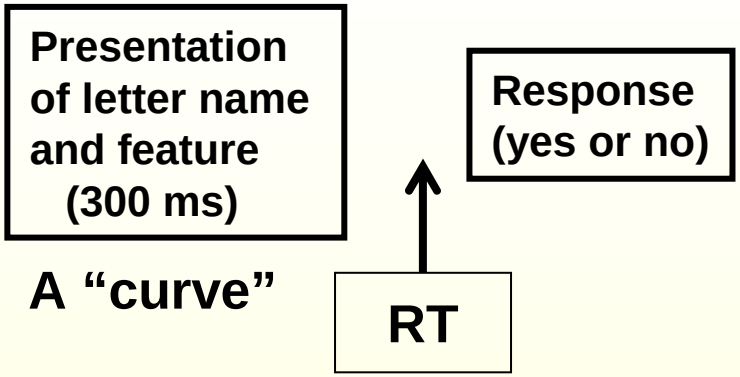
KEY POINTS

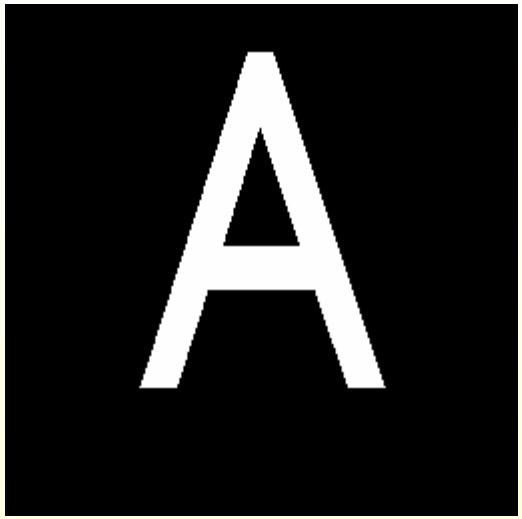
- Central Cognition (Soar)
  - Abstract, symbolic visual representations
  - Domain knowledge (goals, states, task constraints)
  - Controls construction, transformation, generation, and inspection
- Vision / Visual Imagery (SVI)
  - Quantitative and depictive visual representations
  - Leverages mechanisms provided by visual perception.
  - Constructs and generates what it is "told"
  - Provides perceptions based on what it "sees"
  - Enables novel composition of previously perceived objects
  - Reacquires knowledge "abstracted away" during initial perception

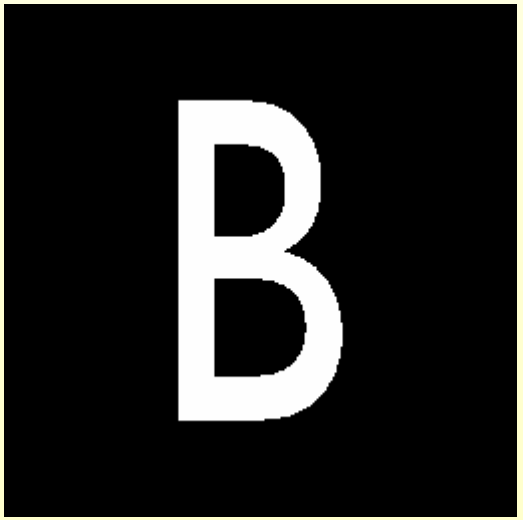
OUTLINE

- REVIEW
- CURRENT ARCHITECTURE (SVI)
- EXPERIMENTAL RESULTS
- FUTURE WORK

DEPICTIVE EXPERIMENT ALPHABET FEATURES



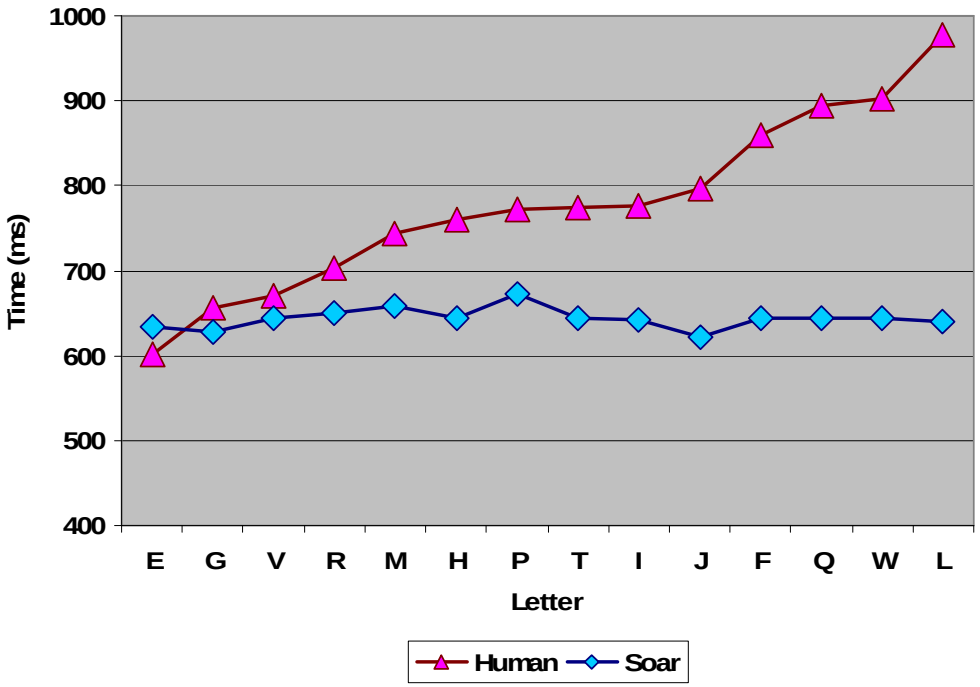


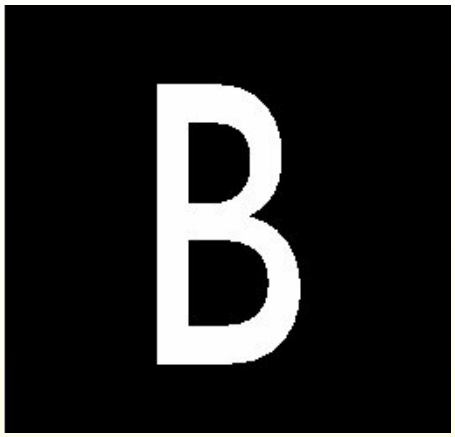


- Emphasized inspection of object features
  - curve
  - symmetry
  - enclosed space
- Depictive (pixel) representations
- Shape (vertices) stored in VisualLTM so had to "construct" visual representation
- External environment, non-visual interaction

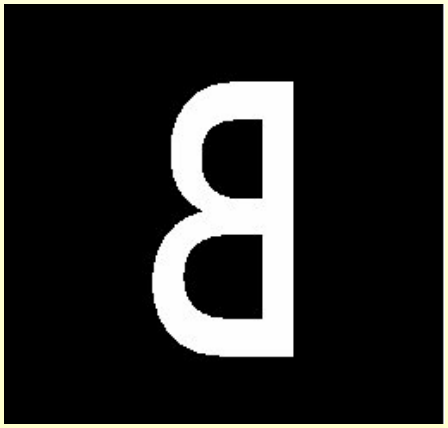
DEPICTIVE EXPERIMENT SYMMETRY

- Transform representation along axis of symmetry
- Make comparison by 'subtracting out" differences
- New capability (+)
- No correlation with human data (-)
  - Transformation in one "cycle".
  - No maintenance of visual representation





Transform



OUTLINE

- REVIEW
- CURRENT ARCHITECTURE (SVI)
- EXPERIMENTAL RESULTS
- FUTURE WORK

EXPERIMENTAL DOMAIN CONSTRAINTS

- Interactive domain and not a “question and answer” task
- Emphasize the interaction with bottom-up visual perception and top-down visual imagery processing to evaluate the perceive-imagine-reperceive cycle
- Evaluate both visual-spatial and visual-depictive imagery
- Exercise major visual imagery functionalities (*construct, transformation, generation, inspection*)

“SCOUT” DOMAIN

- Agent imagines what it and teammate can see (field of view)
or teammate? (Decision)
buildings, bridges)
location and orientation
based on verbal report
verbal reports

LEGEND:
Scout Enemy

NUGGETS & COAL

- NUGGETS
  - Answered some of questions from last year
    - Types of representations
    - Intersection between high-level vision and visual imagery
  - Architectural components are relatively stable
  - Simulation is up and running

- COAL
  - Determination of when to use which representation without a big “switch”
  - Unclear as to details of specific algorithms
  - Processing with concurrent visual perception and visual imagery unknown (resource constraints)