The BICA Cognitive Decathlon: A Test Suite for Biologically-Inspired Cognitive Agents

Shane T. Mueller, Ph.D.
Matt Jones, Ph.D.
Klein Associates Division, ARA Inc.

Brandon S. Minnery, Ph.D.
Julia M. H. Hiland
The MITRE Corporation
BICA Background

Experimental Data    Theoretical Models

Psychology

- Experimental results from psychology and neuroscience
- Insight into neural functions, organization, and processing
- Insight into how humans learn

Neuroscience

- Decomposition into fundamental functional components
- Modeling of component subsystems
- Individual functions and communication among components

New Computational Models of Cognition

- Comprehensive computational models inspired by human cognition and neural mechanisms
- Performance validation through complex “grand challenge” problems
Levels of Modeling

16th BRIMS Conference  26-29 March 2007

BICA Focus

cognition
systems
areas
columns
circuits

1 m
10 cm
1 cm
1 mm
100 μm
1 μm
1 Å

Function Clearer
Structure Less Apparent
Function Clearer
Structure Less Apparent

Wow!

Function Clearer
Structure Less Apparent
Function Clearer
Structure Less Apparent

1 m
10 cm
1 cm
1 mm
100 μm
1 μm
1 Å

CNS
Systems
Maps
Networks
Neurons
Synapses
Molecules
Testing the Goals of BICA

- Embodied, Learning agent with skills of a 2-year-old child
- Turing's Test:
  - A test of behavior, not thought
  - Traditional perspective: important part of intelligence is talking, not acting
  - Embodied cognition rejects the this notion
  - 'Better than human' is 'distinguishible from human'.
Harnad's (2004) Hierarchy of Turing Tests

- Indistinguishable for limited task
- Indistinguishable in verbal context
- Indistinguishable in sensorimotor context
- Indistinguishable in internal structure
- Indistinguishable in physical structure
Design of Tests

- A single test works against a primary goal of BICA: a flexible learning agent
- Multiple small tests work against another goal: end-to-end complete comprehensive agent, produce “Frankenmodel”

Solution:
- Detailed tests of core competencies
- Complex tests requiring integration
- Prediction of brain activations
- Indistinguishable for robust qualitative patterns
BICA Three-part Test

Goals: Biological, embodied, complete learning agent

Challenge Scenarios

Decathlon Events

Cognitive, Perceptual, & Motor Skills

Behavioral & biological comparison
<table>
<thead>
<tr>
<th>Basic Cognitive Skills</th>
<th>Minimum Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Vision</strong></td>
<td><strong>Y1</strong></td>
</tr>
<tr>
<td>Invariant Object ID</td>
<td>75% accuracy for learned objects</td>
</tr>
<tr>
<td>Object ID w Size Discrim</td>
<td>75% accuracy for discriminable objects</td>
</tr>
<tr>
<td>Object ID w Rotation</td>
<td>75% accuracy</td>
</tr>
<tr>
<td>Visual Action / Event Recog</td>
<td>60% accuracy</td>
</tr>
<tr>
<td><strong>Search &amp; Navigation</strong></td>
<td></td>
</tr>
<tr>
<td>Visual Search</td>
<td>75% accuracy; color pop-out &amp; deliberate search functions</td>
</tr>
<tr>
<td>Simple Navigation</td>
<td>Avoid obstacles in cluttered room after 3 learning trials.</td>
</tr>
<tr>
<td>Multi-point Navigation</td>
<td>Solution times approx. linear with target number</td>
</tr>
<tr>
<td>Embodied Search</td>
<td>Find target by searching on average fewer than 75% of locations</td>
</tr>
<tr>
<td>Reinforcement Learning</td>
<td>Exploit maximal reward location (50%+ trials on best payoff)</td>
</tr>
<tr>
<td><strong>Manual Control &amp; Learning</strong></td>
<td></td>
</tr>
<tr>
<td>Motor Mimicry</td>
<td>Success at 25% of actions</td>
</tr>
<tr>
<td>1-Hand Manipulation</td>
<td>Success at 25% of actions</td>
</tr>
<tr>
<td>2-Hand Manip &amp; Construct</td>
<td>Success at 25% of actions</td>
</tr>
<tr>
<td>Device Mimicry</td>
<td>Success at 25% of actions</td>
</tr>
<tr>
<td>Intention Mimicry</td>
<td>Success at 25% of actions</td>
</tr>
<tr>
<td><strong>Knowledge Learning</strong></td>
<td></td>
</tr>
<tr>
<td>Episodic Recog Memory</td>
<td>60% accuracy; strength effects</td>
</tr>
<tr>
<td>Semantic Memory/Categories</td>
<td>25% accurate categorization performance.</td>
</tr>
<tr>
<td><strong>Language / Concept Learning</strong></td>
<td></td>
</tr>
<tr>
<td>Object-Noun Mapping</td>
<td>Success naming on 25% of trials</td>
</tr>
<tr>
<td>Property-Adjective</td>
<td>Success naming on 25% of trials</td>
</tr>
<tr>
<td>Relation-Preposition</td>
<td>Success naming on 25% of trials</td>
</tr>
<tr>
<td>Action-Verb</td>
<td>Success naming on 25% of trials</td>
</tr>
<tr>
<td>Coord Action-Verb</td>
<td>Success naming on 25% of trials</td>
</tr>
<tr>
<td><strong>Simple Motor Control</strong></td>
<td></td>
</tr>
<tr>
<td>Eye Movements</td>
<td>Ability to track smoothly moving objects</td>
</tr>
<tr>
<td>Aimed Manual Move</td>
<td>Ability to point to target.</td>
</tr>
</tbody>
</table>

○ Test with Virtual Robot Embodiment
● Test with Mechanical Robot Embodiment
Basic Cognitive Skills
Basic Skills: Visual Identification

- Visual identification of objects, obstacles, events.
- Learns: component parts, configuration of parts, size, color, etc.

Same or Different?
Basic Cognitive Skills: Search & Navigation

- Tests ability to move through environment and learn where things are.
- Important tests of learning through exploration and reinforcement.

Find the blender.
Basic Skills: Manual Control & Learning

- Tests ability to learn procedural motor control through demonstration.
- Advanced stages involve learning to control novel device or tool to achieve goal.

Hold the blender like this.
Basic Skills: Language

- Language learning *grounded* physical world
- Forms basis of ability to learn tasks through instruction.

This is a silver blender.
Basic Cognitive Skills: Knowledge

- Tests ability to learn facts about events in the world

Have you seen this before?
Basic Cognitive Skills: Simple Motor Control

- Determine whether eye and hand movements are biological

Look at the target.

Touch the target.
Challenge Scenarios

- Complex, end-to-end tests of capability
- Build on core skills tested in decathlon to reduce burden and
- "The Egg Hunt"
- "Show and Tell"
- "Widget Factory"
- Open-ended tasking
Scenario: The Egg Hunt (Find a Hidden Object)

- Agent placed in room, shown an object, and asked to find and retrieve copy in another room. Verbal hints and constraints may be given to agent.

- Skills:
  - Novel object ID, learning by instruction, Navigation, search, meta-tagging.

- Preliminaries Decathlon Events:
  - Simple Navigation, TSP, Object Identification

- Evaluation
  - Success d'
  - Improvement when same search cue is repeated
Scenario: Show and Tell
(Language-object-action integration)

Teacher performs a task and explains its actions while agent watches. Teacher asks agent to do task, giving appropriate feedback. Later, teacher may give verbal instructions to perform a task, or ask agent to describe a previously-learned task agent performs.

- Skills: Language/conceptual mapping, social interaction, tuning with reinforcement, simple construction

- Prelim Events: Language/object mapping, relational mapping, simple construction

- Evaluations: Probability of success at learning skill; ability to generalize, ability to describe learned and novel task
Scenario: Widget Factory Search, Assembly, & Goal Discovery

- Agent is a piece-worker in a widget factory requiring objects to be gathered and assembled for reward. Components appear probabilistically in different rooms.
- Open-ended ecology of rewards requiring agent to discover behaviors and goals that are useful, providing a simplified environment in which tool construction/use is possible.
- Skills: Assembly, search, reinforcement learning, episodic/semantic learning
- Preliminary Decathlon Events
  - Simple Construction, simple search, navigation, foraging/TSP
- Evaluation
  - Increase in earned points, probability of making valuable time/cost trade-offs
Biofidelity Assessment

 Designed by MITRE
 By being faithful to brain organization, we may get beyond traditional AI approaches.
 Teams required to model brain activity at functional, module-level; with commitments to brain regions and interconnectivity
 As program continued, predictions from models about BOLD activity (fMRI) in decathlon tasks would be required.
### Evaluation Milestones

<table>
<thead>
<tr>
<th></th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Egg Hunt</strong></td>
<td>Find 10% of present targets</td>
<td>Search accuracy 20%; learn from 10% of experiences</td>
<td>Search accuracy 40%; learn from 20% of experiences</td>
<td>Search accuracy 60%; learn from 30% of experiences</td>
<td>Search accuracy 80%; learn from 40% of experiences</td>
</tr>
<tr>
<td><strong>Show &amp; Tell</strong></td>
<td></td>
<td>Assembly of simple objects; Learning accuracy 20%</td>
<td>Assembly of novel objects; Learning accuracy 40%</td>
<td>Multistage construction; Learning accuracy 60%</td>
<td>Verbal instruction; Learning accuracy 80%</td>
</tr>
<tr>
<td><strong>Widget Factory</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Decathlon</strong></td>
<td>Demonstrate competency in 3 of 7 tested skills</td>
<td>Demonstrate competency in 7 of 16 tested skills</td>
<td>Demonstrate competency in 15 of 21 tested skills</td>
<td>Demonstrate competency in 23 or 23 tested skills</td>
<td>Demonstrate competency in 23 of 23 tested skills</td>
</tr>
<tr>
<td><strong>Biofidelity</strong></td>
<td>Mapping of model components to brain structures</td>
<td>Comparison of model activity to published brain data</td>
<td>Comparison of model activity to published brain data</td>
<td>Comparison of model activity to fMRI data on BICA tasks</td>
<td>Comparison of model activity to fMRI data on BICA tasks</td>
</tr>
</tbody>
</table>
Conclusions and Post-mortem

- DARPA Phase I spent 15 months & a lot of money in design phase
  - It got a bunch of great ideas
  - Some of them will find new homes
- Phase II has yet to be funded....
- Eurocub robot project
- The test design can inform future efforts