Life-long assurance of Meaningful Human-Autonomy Teaming

- Meaningful Human Control
- Cognitive agents and trust
- Human-Autonomy Team (HAT)
- Meaningful-HAT
- Readiness levels
- Building and Testing
- Conclusions

Ms Kate Yaxley  Dr Keith Joiner, CSC
Artificial intelligence

Narrow AI
  – Task specific

General AI
  – Smart Autonomous System (SAS)
  – Cognitive
  – Allowed to choose
Meaningful Human Control

K1 Predictable, reliable and transparent technology
K2 Accurate information for the user of the outcome sought, operation and function of technology, and the context of use
K3 Timely human action and a potential for timely intervention
K4 Accountability to a certain standard

Cognitive-Cyber Physical Systems

Perform T&E to assure K1, K2

Trust warriors to assure K3, K4

Trust between cognitive agents

- Risk relationship between truster and trustee
- Allows for sharing of ideas, tasks and outcomes
- Can be shaped by environmental, cultural or learned experiences
- Transparency and communication help to build trust between agents
Trusted autonomy

Free to choose a truster and trustee relationship

Must also consider

- Reliability
- Performance
- Error type
- Transparency
- Shared awareness
- Shared intent

Performance appraisal of Human Teams
Human-Autonomy Teaming (HAT)

- Operator interaction
- Human-autonomy communication
- Transparency

Meaningful-Human-Autonomy Teaming (M-HAT)

- Meaningful interaction (learning and reward)
- Human-autonomy communication
- Transparency
- Trust

Technology Readiness Level

<table>
<thead>
<tr>
<th></th>
<th>Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Analytic study</td>
</tr>
<tr>
<td>3</td>
<td>Individual components</td>
</tr>
<tr>
<td>4</td>
<td>Early integration</td>
</tr>
<tr>
<td>5</td>
<td>Laboratory integration</td>
</tr>
<tr>
<td>6</td>
<td>Laboratory or Modelling and Simulation</td>
</tr>
<tr>
<td>7</td>
<td>Demonstration in operational environment</td>
</tr>
<tr>
<td>8</td>
<td>Developmental Test and Evaluation</td>
</tr>
<tr>
<td>9</td>
<td>Operational use or operational testing</td>
</tr>
</tbody>
</table>

## SAS Education Level

<table>
<thead>
<tr>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A(a) Remembering</td>
<td>Recall game-play, recite an instruction</td>
</tr>
<tr>
<td>B(b) Understanding</td>
<td>Explain process using own words, explain movements</td>
</tr>
<tr>
<td>C(c) Applying</td>
<td>Create new moves, create a procedure for others to follow</td>
</tr>
<tr>
<td>D(d) Analysing</td>
<td>Troubleshoot, identify tasks for training, conduct a test</td>
</tr>
<tr>
<td>E(e) Evaluating</td>
<td>Select the most suitable player, justify a process</td>
</tr>
<tr>
<td>F(f) Creating</td>
<td>Integrate information from multiple sources to create a solution, network with others</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TRL</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A (a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B (b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C (c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E (e)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F (f)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SAS Readiness Level (SRL)
SAS Readiness Level

- Allows for the classification of a SAS system in terms of technological and educational readiness

- Ensures SAS used for teams have the technical and cognitive ability to build effective relationships, contribute and learn
AlphaGo

TRL 7 – Demonstration in operational environment

SEL A – Recall game-play

SRL: 7. Abc

SEL b – explain process using own words (machine language to teach AlphaGo Zero)

SEL c – create new moves, yet no procedure to follow
Building M-HAT

- Team building between humans and technology
- Reward trustworthiness
- Transparency in actions and communication
Testing SAS

- Complete in phases
  - Simulation
  - HWIL
  - Operating environments
- Iterative education
- Iterative technology development
- Iterative training

Test Team

- Supervisor
  - As per HRI
  - Human agent

- WAI Agent
  - Performs same role as supervisor
  - SAS, but not M-HAT
  - Interrogates and assesses

- Educator
  - Assists with assessing SEL

- Testing Team responsible for
  - K1
  - K2
  - K3
  - K4
• Operator
  – Software assurance

• Mechanic
  – Hardware assurance

• SAS Trainers responsible for assuring
  – K1
  – K2
  – K4
WAI Agent
  – K1
  – K2
Peer
  – K3
  – K4
Bystander
  – K3
  – K4
Conclusions

■ Teaming will exist both at the ground and supervision level

■ HRI concept to assure MHC

■ Requires understanding of both technological and cognitive levels to assure M-HAT

■ Using SRL, M-HAT and MHC, situational awareness of teams and team dynamics may be developed
QUESTIONS
<table>
<thead>
<tr>
<th></th>
<th>Known to SAS</th>
<th>Not Known to SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Known to</strong> WAI &amp; M-HAT</td>
<td>Open area/Arena</td>
<td>Blind spot</td>
</tr>
<tr>
<td></td>
<td>Low Risk</td>
<td>Low-Moderate Risk</td>
</tr>
<tr>
<td><strong>Not known to</strong> WAI &amp; M-HAT</td>
<td>Hidden area/Façade</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>Moderate risk</td>
<td>High Risk</td>
</tr>
</tbody>
</table>
Beyond Phase 1

Situation Awareness

SRL

Phase 1

Phase 2

Phase 3

Beyond