ACT-R Workshop Schedule

Opening: ACT-R from CMU’s Perspective
  9:00 - 9:45  Overview of ACT-R -- John R. Anderson
  9:45 – 10:30 Details of ACT-R 6.0 -- Dan Bothell

Break: 10:30 – 11:00

Presentations 1: Architecture
  11:00 – 11:30 Functional constraints on architectural mechanisms -- Christian Lebiere
  11:30 – 12:00 Retrieval by Accumulating Evidence in ACT-R -- Leendert van Maanen
  12:00 – 12:30 A mechanism for decisions in the absence of prior reward -- Vladislav D. Veksler

Lunch: 12:30 – 1:30

Presentations 2: Extensions
  1:30 – 2:00 ACT-R forays into the semantic web -- Lael J. Schooler
  2:00 – 2:30 Making Models Tired: A Module for Fatigue -- Glenn F. Gunzelmann
  2:30 – 3:00 Acting outside the box: Truly embodied ACT-R -- Anthony Harrison
  3:00 - 3:30 Interfacing ACT-R with different types of environments and with different techniques: Issues and Suggestions.-- Michael J. Schoelles

Break: 3:30 – 4:00

Panel: 4:00 – 5:30: Future of ACT-R from a non-CMU Perspective
  Danilo Fum, Kevin A. Gluck, Wayne D. Gray, Niels A. Taatgen, J. Gregory Trafton, Richard M. Young
ACT-R/GPD

Vladislav “Dan” Veksler
Purpose

• Propose a mechanism in ACT-R for making decision in the absence of prior reward

• Not meant to replace the current ACT-R reward-based decision mechanism, but rather to complement it
Current ACT-R model of choice
Current ACT-R model of choice

Central Executive/Procedural Module
Current ACT-R model of choice

Central Executive/Procedural Module

based on prior reward
(Reinforcement Learning; RL)
Current ACT-R model of choice

- ACT-R model of human choice is based on Reinforcement Learning (RL)
  - a formal model of human/animal trial-and-error behavior
  - predicts human choice based on prior reward/punishment
  - psychological and biological evidence (e.g. Holroyd & Coles, 2002)

- However, much of human choice is based on other information
The 2-goal problem

- An agent is tasked with achieving some goal, A
- Then the agent is tasked with achieving B, in the same environment

- RL would perform on the 2nd task no better than one the 1st
- Humans learn their environment while achieving A, thus helping to reduce their time to achieve B
The 2-goal problem
The 2-goal problem
The 2-goal problem

when goal="key" goto left-arm

when goal="key" goto right-arm
The 2-goal problem

when goal="key"
goto left-arm

when goal="key"
goto right-arm
The 2-goal problem

when goal="key"
goto left-arm

when goal="key"
goto right-arm
The 2-goal problem

- when goal="key" +
  goto left-arm

- when goal="key" -
  goto right-arm
The 2-goal problem

when goal="key" +
goto left-arm

when goal="key" -
goto right-arm
The 2-goal problem

when goal="key"
  goto left-arm +
when goal="key"
  goto right-arm −
The 2-goal problem

when goal="key" +
goto left-arm

when goal="key" -
goto right-arm
The 2-goal problem

when goal="key"
+ goto left-arm
when goal="key"
− goto right-arm
The 2-goal problem

when goal="key" +
go to left-arm

when goal="key" -
go to right-arm

when goal="purse"
go to left-arm

when goal="purse"
go to right-arm
The 2-goal problem

when goal="key" +
goto left-arm

when goal="purse"
goto left-arm

when goal="key" -
goto right-arm

when goal="purse"
goto right-arm
The 2-goal problem

when goal="key" +
goto left-arm

when goal="key" -
goto right-arm

when goal="purse" +
goto left-arm

when goal="purse" -
goto right-arm
The 2-goal problem

humans make the correct choice >50% of the time (Stevenson, 1954; Quartermain & Scott, 1960)
ACT-R can use declarative information in decisions

• An ACT-R model can be written to perform this task
  
  • e.g. storing all attended items as declarative chunks:
    [:location left :item key]
    [:location left :item purse]
    [:location right :item candy]
    ...

• However, there are no architectural constraints for doing this
  
  • no system-level prediction for how humans make decisions in the absence of reward
SNIF-ACT

- Pirolli & Fu (2003) ACT-R model of web-browsing (SNIF-ACT)
  - in web-browsing new links are encountered with no prior reward
  - choose-link production utilities based on associative knowledge
SNIF-ACT

- Piroli & Fu (2003) ACT-R model of web-browsing (SNIF-ACT)

- in web-browsing new links are encountered with no prior reward

- choose-link production utilities based on associative knowledge
SNIF-ACT

- Pirolli & Fu (2003) ACT-R model of web-browsing (SNIF-ACT)
  - in web-browsing new links are encountered with no prior reward
  - choose-link production utilities based on associative knowledge
  - limited to web-browsing type tasks
  - no associative learning *
    - associative knowledge comes from PMI engine (Pointwise Mutual Information; Turney, 2001)
    - PMI predicts the strength of association between words based on co-occurrence
Voicu & Schmajuk

- Voicu & Schmajuk (2002) model of navigation
  - similar to SNIF-ACT, decisions based on spreading activation from the goal
  - simulates qualitative effects of latent learning, shortcut, detour behavior
  - limited to single-goal navigation tasks
Goal-Proximity Decision Mechanism (GPD)

- Utility of choice is predicted based on its associative strength to current goal
  - inherent value of the goal spreads to options

  ➤ as a complement to the RL mechanism in ACT-R
Goal-Proximity Decision Mechanism (GPD)

- Given goal G, and a choosing between options A and B
  - retrieve A or B from memory
  - option with higher association strength to G more likely to be retrieved
  - association strengths reflect experienced item proximity
Goal-Proximity Decision Mechanism (GPD)

- Given goal G, and a choosing between options A and B
  - retrieve A or B from memory
  - option with higher association strength to G more likely to be retrieved
  - association strengths reflect experienced item proximity
Goal-Proximity Decision Mechanism (GPD)

- Given goal G, and a choosing between options A and B
  - **retrieve** A or B from memory
  - option with higher *association strength* to G more likely to be retrieved
  - association strengths reflect experienced item *proximity*

... X A J B C G ...
Implementing GPD in ACT-R
Implementing GPD in ACT-R

- Given goal G, and a choosing between options A and B

  - retrieve A or B from memory

  - option with higher association strength to G more likely to be retrieved

  - association strengths reflect experienced item proximity
Implementing GPD in ACT-R

- Given goal G, and a choosing between options A and B
  - retrieve A or B from memory
  - option with higher association strength to G more likely to be retrieved
  - association strengths reflect experienced item proximity
Implementing GPD in ACT-R

- Given goal G, and a choosing between options A and B
  - \textbf{retrieve} A or B from memory
  - +retrieval>
  - option with higher \textit{association strength} to G more likely to be retrieved
  - association strengths reflect experienced item \textit{proximity}
Implementing GPD in ACT-R

• Given goal G, and a choosing between options A and B
  
  • **retrieve** A or B from memory

  • option with higher **association strength** to G more likely to be retrieved

  • association strengths reflect experienced item **proximity**

* may be better to implement this at system level (goal module?)
Implementing GPD in ACT-R

• Given goal G, and a choosing between options A and B

  • **retrieve** A or B from memory

  +retrieval>

  • option with higher **association strength** to G more likely to be retrieved

  • association strengths reflect experienced item **proximity**
Implementing GPD in ACT-R

• Given goal G, and a choosing between options A and B

  • retrieve A or B from memory

  • option with higher association strength to G more likely to be retrieved

  • association strengths reflect experienced item **proximity**

    • episodic buffer (list of recently attended chunks)

    • association strength between two chunks is incremented proportional to their proximity in the episodic buffer (in error-driven fashion)
Associative Learning

• given a new episode, $j$
  • for each episode in episodic buffer, $i$
    • decrease activation of $i$, $a_i$, by $\vartheta$
    • increase $S_{ji}$

$$\Delta S_{ji}(n) = \beta [a_i(n) - S_{ji}(n-1)]$$

• push episode $j$ into episodic buffer
• $a_j = 1$

- $S_{ji}(n)$ – strength of association between $j$ and $i$ at time $n$
- $\Delta S_{ji}(n)$ – change in $S_{ji}$ at time $n$
- $a_i$ – activation of $i$ at time $n$
- $\beta$ – learning rate
- $\vartheta$ – activation decay
Associative Learning

• given a new episode, $j$
  • for each episode in episodic buffer, $i$
    • decrease activation of $i$, $a_i$, by $\vartheta$
    • increase $S_{ji}$

$$\Delta S_{ji}(n) = \beta [a_i(n) - S_{ji}(n-1)]$$

• push episode $j$ into episodic buffer
• $a_j = 1$

Episodic Memory

- $S_{ji}(n)$ – strength of association between $j$ and $i$ at time $n$
- $\Delta S_{ji}(n)$ – change in $S_{ji}$ at time $n$
- $a_i$ – activation of $i$ at time $n$
- $\beta$ – learning rate
- $\vartheta$ – activation decay
Associative Learning

• given a new episode, $j$
  • for each episode in episodic buffer, $i$
    • decrease activation of $i$, $a_i$, by $\varrho$
    • increase $S_{ji}$

$$\Delta S_{ji}(n) = \beta [a_i(n) - S_{ji}(n-1)]$$

• push episode $j$ into episodic buffer
• $a_j = 1$

Episodic Memory

$\text{FIFO queue}$

• $S_{ji}(n)$ – strength of association between $j$ and $i$ at time $n$
• $\Delta S_{ji}(n)$ – change in $S_{ji}$ at time $n$
• $a_i$ – activation of $i$ at time $n$
• $\beta$ – learning rate
• $\varrho$ – activation decay
Associative Learning

- given a new episode, \( j \)
  - for each episode in episodic buffer, \( i \)
    - decrease activation of \( i, a_i \), by \( \varphi \)
    - increase \( S_{ji} \)

\[
\Delta S_{ji}(n) = \beta [a_i(n) - S_{ji}(n-1)]
\]

- push episode \( j \) into episodic buffer
- \( a_j = 1 \)

Episodic Memory

- \( S_{ji}(n) \) – strength of association between \( j \) and \( i \) at time \( n \)
- \( \Delta S_{ji}(n) \) – change in \( S_{ji} \) at time \( n \)
- \( a_i \) – activation of \( i \) at time \( n \)
- \( \beta \) – learning rate
- \( \varphi \) – activation decay

* may be better to use ACT-R activation (which includes decay)
Associative Learning

- given a new episode, $j$
  - for each episode in episodic buffer, $i$
    - decrease activation of $i$, $a_i$, by $\varepsilon$
    - increase $S_{ji}$
      - push episode $j$ into episodic buffer
    - $a_j = 1$

**Associative Learning**

$$\Delta S_{ji}(n) = \beta [a_i(n) - S_{ji}(n-1)]$$

**Reinforcement Learning**

$$\Delta U_i(n) = \alpha [R_i(n) - U_i(n-1)]$$

and $i$ at time $n$

- $\Delta S_{ji}(n)$ – change in $S_{ji}$ at time $n$
- $a_i$ – activation of $i$ at time $n$
- $\beta$ – learning rate
- $\varepsilon$ – activation decay
Experiment

- Friday at 14:00
RMSE for 2-choice (left) and 3-choice (right) mazes

- Random: RMSE = 39.70
- RL: RMSE = 21.91
- GPD: RMSE = 3.16
- Ideal Performer: RMSE = 6.21

- Random: RMSE = 45.79
- RL: RMSE = 18.29
- GPD: RMSE = 7.95
- Ideal Performer: RMSE = 16.34
GPD versus RL

• GPD is not meant to replace RL
  • it is obvious that much of human choice is based on reward/punishment

• GPD is meant to be a complement to RL

• How GPD and RL interact is a topic for future research
Second Life Simulations

257.492 VISION SET-BUFFER-CHUNK VISUAL-LOCATION VISUAL-LIGHT
257.542 PROCEDURAL PRODUCTION-FIRED DO-ALL-SEE
257.627 VISION SET-BUFFER-CHUNK VISUAL BTM498
257.677 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-ATTEND
257.727 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-CHECKINDEX
257.727 PROCEDURAL PRODUCTION-FIRED COMPARE-RETRIEVE
257.777 DECLARATIVE SET-BUFFER-CHUNK RETRIEVAL __victorion...
257.822 PROCEDURAL PRODUCTION-FIRED COMPARE-SETBEST

257.877 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEN...
257.877 VISION SET-BUFFER-CHUNK VISUAL-LOCATION VISUAL-
257.927 PROCEDURAL PRODUCTION-FIRED DO-ALL-SEE
258.012 VISION SET-BUFFER-CHUNK VISUAL BTM498
258.062 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-ATTEND
258.112 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-CHECKINDEX...
258.162 PROCEDURAL PRODUCTION-FIRED COMPARE-RETRIEVE
258.162 DECLARATIVE SET-BUFFER-CHUNK RETRIEVAL __Object...
258.212 PROCEDURAL PRODUCTION-FIRED COMPARE-SETBEST

258.252 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEn...
258.312 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-DONE
258.362 PROCEDURAL PRODUCTION-FIRED DO-BTN-LOOK
258.362 VISION SET-BUFFER-CHUNK VISUAL-LOCATION VISUAL-
258.412 PROCEDURAL PRODUCTION-FIRED DO-ALL-SEE
258.497 VISION SET-BUFFER-CHUNK VISUAL BTM498
258.547 PROCEDURAL PRODUCTION-FIRED DO-BTN-ATTEND
258.557 MOTOR MOVE-CURSOR OBJECT NIL LOC VISUAL-LOCATION

Model moved mouse to (x:40 y:10)
259.029 PROCEDURAL PRODUCTION-FIRED DO-BTN-CLICK
259.029 MOTOR CLICK-MOUSE
259.159 PROCEDURAL PRODUCTION-FIRED DO-BTN-CLICKEDBEST
259.129 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEn...
259.179 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEn...
259.229 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEn...

Model clicked Object
(-agent act "M:5cbacc4-35cf-1e5d-9bbcc-e710e3ff5748")

Perception: 1228703947
Re-attempting perception: 1228703947

Re-attempting perception: 1228703954
259.239 VISION SET-BUFFER-CHUNK VISUAL-LOCATION VISUAL-
259.279 PROCEDURAL PRODUCTION-FIRED COMPARE-BTN-LOCK4UNATTEn...
Second Life Simulations
Second Life

• GPD may perform better than RL in early stages (prior to reward), but...
  • associations between all the objects become confusing after enough exploration
  • and RL eventually outperforms GPD

✶ Caveat:
  • the results may be specific to the given task
  • selective attention and further parameter space exploration may help GPD
ACT-R/GPD

• How GPD may be integrated with the current ACT-R decision mechanism:
  • GPD may be an appropriate mechanism prior to reward,
  • but once there is reward, RL may take over

• In other words, GPD may be useful for approximating what the reward value would be, before actually experiencing that reward value
Autonomy

- GPD model was not altered between tasks
Future Research

• How GPD and RL may interact

• GPD implementation in ACT-R module

• Using ACT-R activation equation in place of episodic activation

• Associative learning implementation in ACT-R module
Questions?