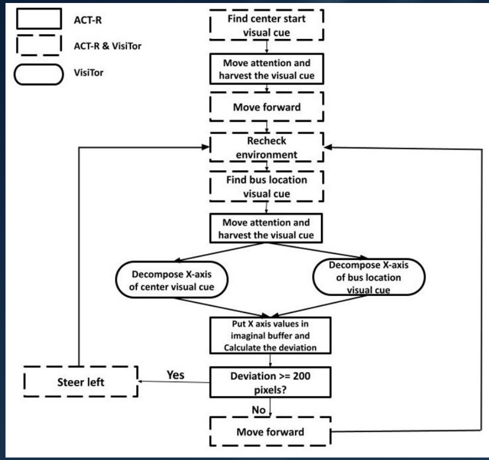
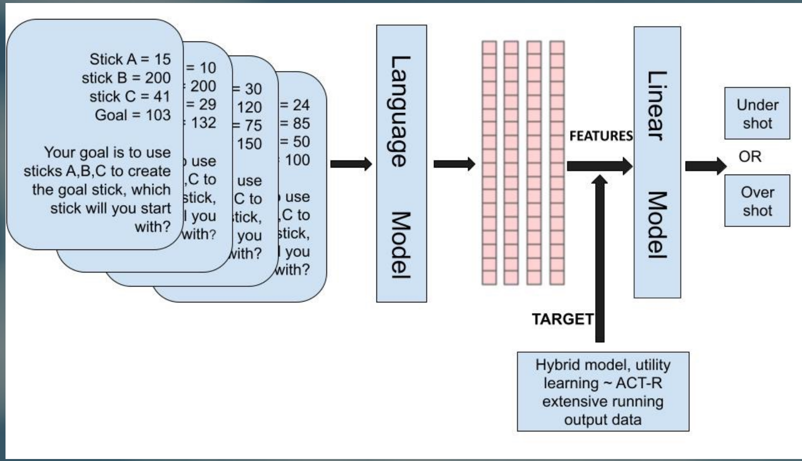
Traversing the Neural-Symbolic Architecture Landscape: From Perceptual Motor Models to Decision-Making with ACT-R – Leading to a Research Proposal

Intern talk for Carnegie Bosch Institution, Jan 11th, 2024

Siyu Wu, MS in Instructional Technology, BS in Applied Finance

First Year PhD in Applied Cognitive Science Lab, College of Information Sciences and Technology, PSU





References

Oltramari, A. 2023. Enabling High-Level Machine Reasoning with Cognitive Neuro-Symbolic Systems. In Proceedings of the *2023 AAAI Fall Symposium on Integrating Cognitive Architectures and Generative Models*. AAAI Press.

Binz, M., & Schulz, E. Turning large language models into cognitive models, arXiv:2306.03917

Ritter, F. E., Tehranchi, F., & Oury, J. D. (2019). ACT-R: A cognitive architecture for modeling cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(3), Paper e1488.

Bagherzadeh, A., & Tehranchi, F. (2022). Comparing cognitive, cognitive instance-based, and reinforcement learning models in an interactive task. Proceedings of ICCM*, The 20th International Conference on Cognitive Modeling.* 1-7.

Wu, S., Ferreira, R., Ritter, F., Walter., L. 2023. Comparing LLMs for Prompt-Enhanced ACT-R and Soar Model Development: A Case Study in Cognitive Simulation. *In Proceedings of the 2023 AAAI Fall Symposium on Integrating Cognitive Architectures and Generative Models*. AAAI Press.

Wu, S., Bagherzadeh, A., Ritter, F., Tehranchi, F. (in press, 2023) Long Road Ahead: Lessons Learned from the (soon to be) Longest Running Cognitive Model. *In Proceedings of 21st International Conference on Cognitive Modeling (ICCM) at the University of Amsterdam*, the Netherlands

Wu, S., Bagherzadeh, A., Ritter, F., Tehranchi, F. (2023, Sep) Cognition Models Bake-off: Lessons Learned from Creating Long-Running Cognitive Models. *Poster in proceedings 16th International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation (SBP-BRIMs)*

Cover Page

Motivation: I aim to revolutionize decision-making processes through the integration of neuro-symbolic architectures and state-of-the-art data-driven approaches.

Strengths: Revived and transformed a stagnant ACT-R project, creating two advanced neuro-symbolic models that have heavy perceptual motor components, leading to three publications.

Technical Expertise:

- Developed two innovative ACT-R models.
- Proficiency in Python for data analytics.

Vision: Fusing neuro symbolic architecture (ACT-R) with language models to improve decision making in domain Q&A tasks.

In this presentation: Tracing my research voyage: *ICCM 2023, BRIMS 2023, AAAI 2023*, and the NEXT PROPOSAL

Long Road Ahead: Lessons Learned from the (soon to be) Longest Running Cognitive Model

Siyu Wu, Amir Bagherzadeh, Frank E. Ritter,
and Farnaz Tehranchi
(sfw5621@psu.edu , abb6024@psu.edu , frank.ritter@psu.edu , farnaz.tehranchi@psu.edu)

7/18/2023
ICCM 2023

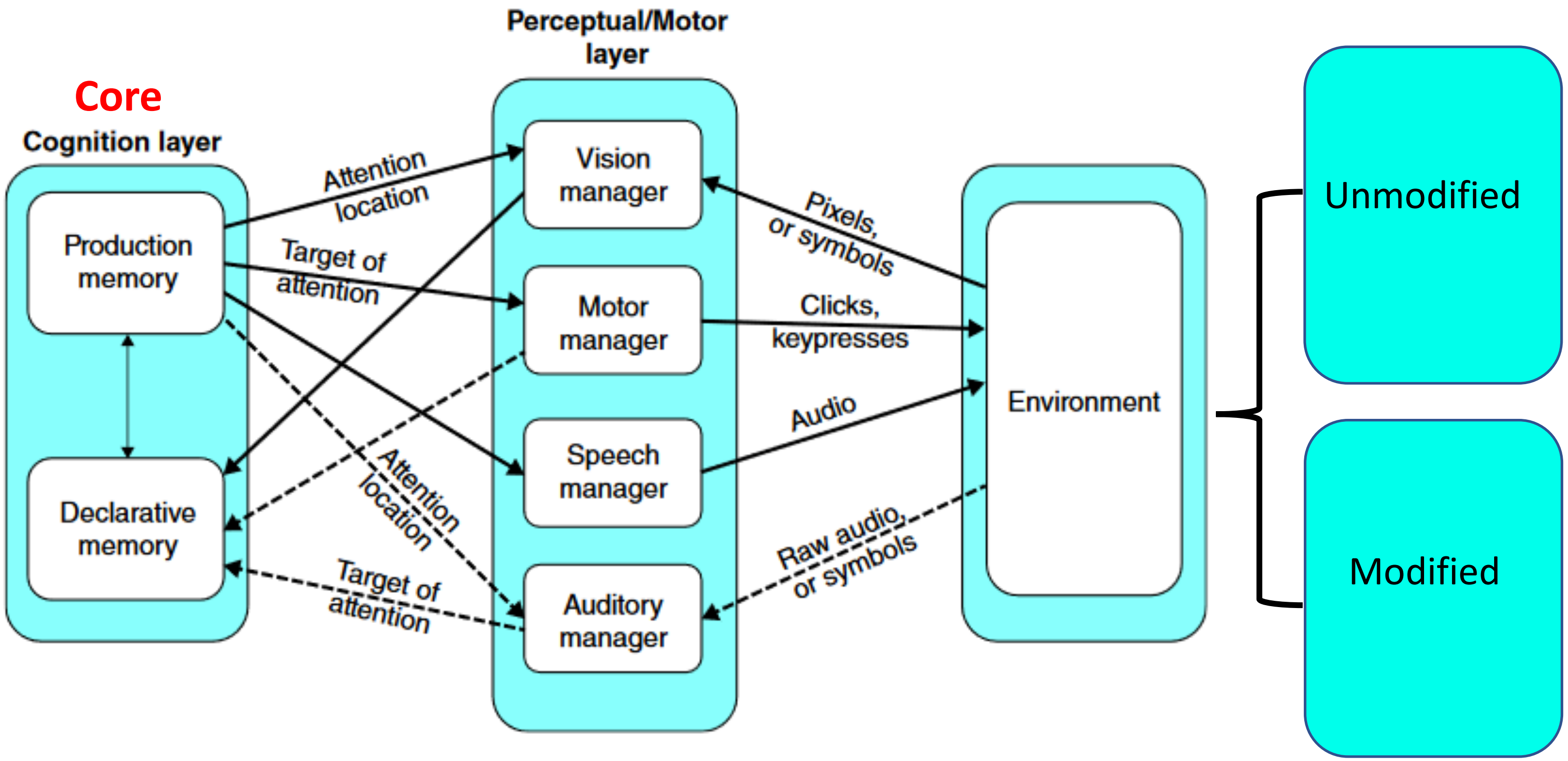
Image adapted from: Ritter, F. E., Tehranchi, F., & Oury, J. D. (2019). ACT-R: A cognitive architecture for modeling cognition. Wiley Interdisciplinary Reviews: Cognitive Science, 10(3), Paper e1488. http://acs.ist.psu.edu/papers/ritterTOip.pdf

ACT-R

- Cognitive architecture
- Modules to implement the fixed mechanisms of cognition

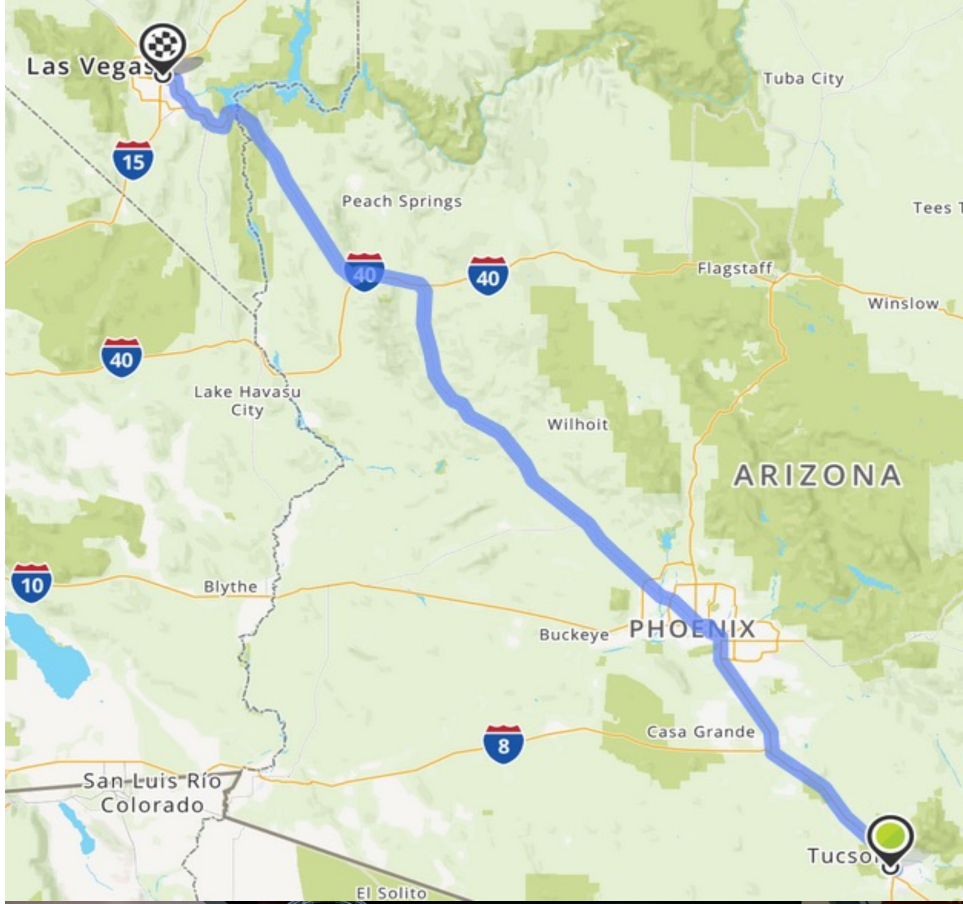
Productions,
if-then rules,
actionable memory

Chunks,
factual memory,
slot-value pairs



Task

- Desert Bus
- Drive from Tucson, AZ to Las Vegas, NV
- Can go a maximum of 45 mph
- ”w” to go forward, “a” to go left
- **Game runs in real-time**
    - No pausing, bus pulls right and slows down



Original

Free remake by
Dinosaur Games
URL here

The Challenge with Previous Driving Simulation: Unmodified Interface

Table 1. Limitations of the Schwartz et al. (2020) model
Did not start the simulation
Did not drive for more than 20 min
Only did drive for short distance mm(s)

Schwartz, D. M., Tehranchi, F., & Ritter, F. E. (2020). Drive the bus: Extending JSegMan to drive a virtual long-range bus. In *Proceedings of ICCM-2020-18th International Conference on Cognitive Modeling* (pp. 1-6).

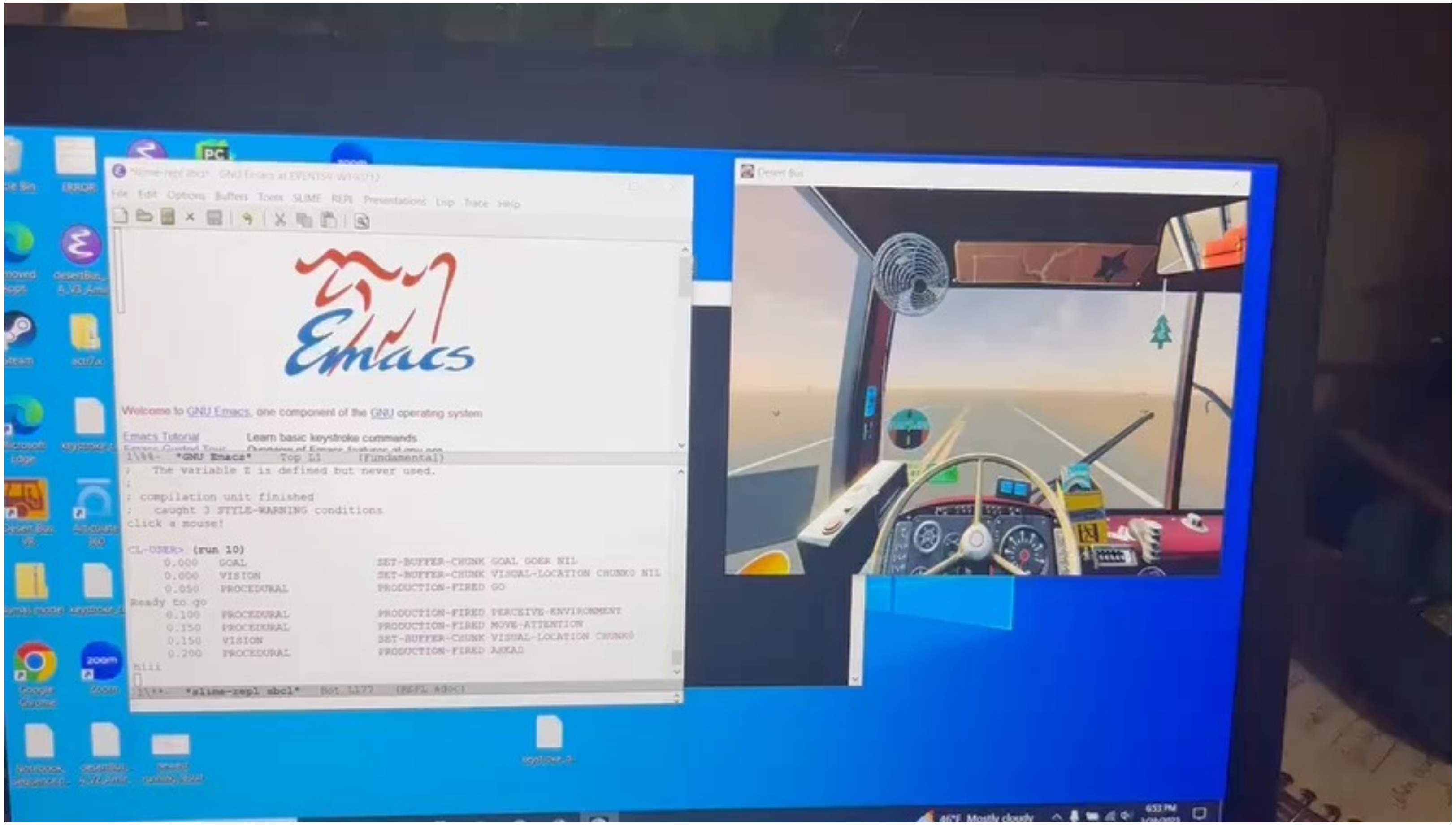
We Created a Better Model

- Achieved a 1200% improvement in longer driving time compared to Schwartz et al.'s (2020) bus driving model, with a maximum running time extended to 4 hours with an average running time of 1 hour
- Incorporated intelligent systems with cognitive modeling techniques (ACT-R)
- Extended capabilities with robotic hands and eyes using VisiTor
- Utilized different declarative chunks and production rules
- Optimized the approach for creating cognitive intelligent agents

Architecture of Interaction

- **ACT-R 7 model with visual and motor capabilities**: The model incorporates a Perceptual-Motor module for accessing interfaces in Macintosh Common Lisp (MCL) and refining behavior through module modifications.
- **VisiTor for interaction with uninstrumented interfaces**: VisiTor is used as a solution to enable ACT-R 7 to interact with uninstrumented interfaces, providing simulated visual attention, mouse, and keyboard actions.
- **Application in Desert Bus task**: In the Desert Bus task, ACT-R directs VisiTor to scan the screen, start the program, accelerate, and steer to keep the bus on the road.
- **VisiTor underwent minor extensions** to simplify object descriptions, incorporate various objects, and facilitate the transfer of motor commands with adjustable durations. The new "longpresskey" feature lets it simulate continuous key presses.

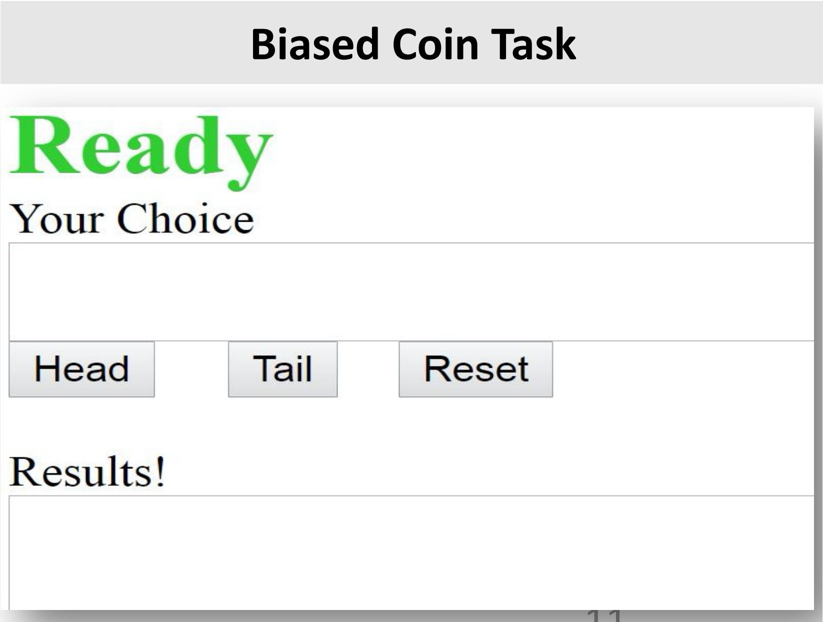
Performance Demonstration: One Loop Execution



VisiTor

- VisiTor is a Python software package available on GitHub that provides simulated hands and eyes for computer interactions.
- It offers two types of functions: motor functions and visual functions. Motor functions simulate mouse clicks, keypresses, and mouse movements to specific screen locations or visual patterns.
- Visual functions in VisiTor include checking for specific visual patterns on the screen, locating patterns within defined modules, and retrieving the current mouse location.
- Previous models have utilized VisiTor for interaction history, such as the "Heads and Tails" model.

Bagherzadeh, A., & Tehranchi, F. (2022). Comparing cognitive, cognitive instance-based, and reinforcement learning models in an interactive task. *Proceedings of ICCM, The 20th International Conference on Cognitive Modeling*. 1-7.



Model

The model has two types of chunks: "drive" and "encoding", with a total of 12 declarative memories.

- The "drive" chunk has slots for "strategy" and "state", with parameters for object items.
- The "encoding" chunk has slots for screen-x locations of visual cues and a deviation slot.

The model utilizes an explicit goal state to control its actions and consists of 13 production rules

- The model investigates the simulation environment to locate and collect the necessary visual cue to initiate gameplay
- It uses the manual buffer to maintain forward motion by holding down the key.
- The model continuously evaluates the environment to gather visual cues related to the bus's location.
- It calculates the difference between the x-axis values of the center line and bus location
- If the difference exceeds 200 pixels, the model turns the bus left by holding down the key; otherwise, it continues moving forward.

Highlights from the Simulation Results

- Implements a superior control mechanism to enhance long-term bus driving performance
- Enhances coordination between perceptual and motor behavior by integrating VisiTor into ACT-R 7, reducing processing time of identifying images to steering from 6.01s to 0.9s
- Leverages VisiTor's extensibility to create new functions, such as the long key press function, to improve task performance.

Cognition Models Bake-off: Lessons Learned from Creating Long-Running Cognitive Models

Siyu Wu, Amir Bagherzadeh, Frank E. Ritter, and Farnaz Tehranchi

(sfw5621@psu.edu , abb6024@psu.edu , frank.ritter@psu.edu , farnaz.tehranchi@psu.edu)

09/2023

Presented at *BRIMS* 2023

What’s New
When Compared to ICCM

- Developed two AI models that can play computer games like humans, interacting with the game through the original interface.
- These models represent different types of human behavior observed in gameplay and incorporate knowledge of the virtual environment.
- Two different design approaches were used for these models, employing different knowledge representations and actions.

Why Important

The comparative analysis of two models on the same simulation task contribute to the development of more efficient and realistic ACT-R agents for various applications beyond gaming.

PennState
UNIFESP
Universidade Federal de São Paulo

Comparing LLMs for Prompt-Enhanced ACT-R and Soar Model Development: A Case Study in Cognitive Simulation

Siyu Wu (sfw5621@psu.edu), Rodrigo F. Souza (rodrigo.ferreira@unifesp.br),
Frank E. Ritter (frank.ritter@psu.edu), Walter T. Lima Jr. (walter.lima@unifesp.br)

AAAI FSS

Oct. 25th, 2023

Nuggets

- Experiments show that large language models (LLMs) have the potential to be used as interactive interfaces to develop ACT-R and Soar models.
- We documented and resolved the mistakes that LLMs made during this integration.
- We also presented a framework of prompt patterns that maximizes LLMs' interaction with artificial cognitive architectures.

Simulation task

- Autonomous driving game on Stream
- Desert Bus

📘 Wu, S., Bagherzadeh, A., Ritter, F. E., & Tehranchi, F. (2023). Cognition models bake-off: Lessons learned from creating long-running cognitive models. In press *BRIMS 23*.

Research questions

How effectively can ChatGPT4 and Google Bard, as Large Language Models (LLMs), serve as interactive interfaces for developing ACT-R and Soar models in the context of cognitive task simulation?

Prompt design

- Providing context
- Posing perception-related queries
- Simulating decision-making scenarios
- Evaluating the system's responses
- Iterative fine tuning

Model development

- Provide Context and Prompt for Declarative Chunk(s)
- Human In the Loop and Query Initial Rule
- Test and Debug the Code
- Human In the Loop and Query Following Rule
- Model Evaluation

“When using ChatGPT to generate production rules in ACT-R, human-in-the-loop combined with synchronous domain-specific training can increase the syntactic quality of the code.”

Evaluate the generated model

```
CL-USER> (run 1)
     0.000   GOAL                   SET-BUFFER-CHUNK GOAL GOER NIL
     0.000   VISION                 SET-BUFFER-CHUNK VISUAL-LOCATION CHUNK0 NIL
     0.000   VISION                 visicon-update
     0.000   PROCEDURAL             CONFLICT-RESOLUTION
     0.050   PROCEDURAL             PRODUCTION-FIRED READY
     0.050   PROCEDURAL             CLEAR-BUFFER VISUAL-LOCATION
     0.050   VISION                 Find-location
     0.050   VISION                 SET-BUFFER-CHUNK VISUAL-LOCATION CHUNK0
     0.050   PROCEDURAL             CONFLICT-RESOLUTION
     0.100   PROCEDURAL             PRODUCTION-FIRED LOOK-AROUND
     0.100   PROCEDURAL             CLEAR-BUFFER VISUAL-LOCATION
     0.100   PROCEDURAL             CLEAR-BUFFER VISUAL
     0.100   PROCEDURAL             CONFLICT-RESOLUTION
     0.185   VISION                 Encoding-complete CHUNK0-1 NIL
     0.185   VISION                 SET-BUFFER-CHUNK VISUAL CHUNK1
     0.185   PROCEDURAL             CONFLICT-RESOLUTION
     0.235   PROCEDURAL             PRODUCTION-FIRED PRESS-W-KEY
Pressing W key
```

("W" accelerates the bus)

Evaluate the generated model

- The prompts that we fed into Chat-GPT 4 generated a model that performs the corresponding behavior. For the DMs, the model has the necessary chunk-types and slots to decide the driving state based on the visual cues it perceives. For the production rules, it sequentially fires get ready, then looks around, sees the visual pattern using the visicon, and then uses the manual buffer to press the key. This model has the potential to interact with the unmodified and novel simulation environment, and might be helpful to some modelers.

- Similarly, Tehranchi & Bagherzadehkhorasani (@PSU but presented award-winning HFES paper), are using LLM to extend Simulated eyes and hands (VisiTor), e.g., for Desert Bus and general vision models (text -> task chunks:

    R1: NL task text -> VisiTor;
    R2: -> OCR;
    R3: -> chatGPT;
    R4: -> DMs;
    Herbal-like OR chatGPT carries on with rules to use chunks

- It is worth noting that the initial code was not correct enough to run on its own. This can be contrasted with success stories told about working Java and Visual Basic, where existing programs may be used. However, in this case, the semantics of Soar and ACT-R are more complicated, and there may not be enough worked examples that were used to create these LLMs.

Bagherzadeh, A., & Tehranchi, F. (2023 in press). Automatic Error Model (AEM) for user interface design: A new approach to include errors and error corrections in a cognitive user model. In *Proceedings of Human Factors and Ergonomics Society*. Outstanding student paper.

Paik, J., Kim, J. W., Ritter, F. E., & Reitter, D. (2015). Predicting user performance and learning in human-computer interaction with the Herbal compiler. *ACM Transactions on Computer-Human Interaction, 22*(5), Article 25

Next Proposal

- Oltramari, A. 2023. Enabling High-Level Machine Reasoning with Cognitive Neuro-Symbolic Systems. In Proceedings of the 2023 AAAI Fall Symposium on Integrating Cognitive Architectures and Generative Models. AAAI Press.
- Binz, M., & Schulz, E. Turning large language models into cognitive models, arXiv:2306.03917
- Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem-solving: Combined influences on operator selection. Cognitive Psychology, 31, 168-217
- Wu, S., Ferreira, R., Ritter, F., Walter., L. 2023. Comparing LLMs for Prompt-Enhanced ACT-R and Soar Model Development: A Case Study in Cognitive Simulation. *In Proceedings of the 2023 AAAI Fall Symposium on Integrating Cognitive Architectures and Generative Models*. AAAI Press.
- Wu, S., Bagherzadeh, A., Ritter, F., Tehranchi, F. (in press, 2023) Long Road Ahead: Lessons Learned from the (soon to be) Longest Running Cognitive Model. *In Proceedings of 21st International Conference on Cognitive Modeling (ICCM) at the University of Amsterdam*, the Netherlands

High Level View

This research aims to enhance language model reasoning by fine-tuning the last layer of embedding of a problem-solving prompt within LLaMA using ACT-R output from a domain-specific task, specifically the building-sticks problem solving by Lovett (1996).

It will deepen our understanding and adoption of neural-symbolic models by demonstrating how cognitive architectures can enhance AI's ability to process and respond to complex, domain-specific queries

Outline

Title: Enhancing Language Models with Neuro-Symbolic Architecture (ACT-R) for decision making / reasoning in Specific Domains

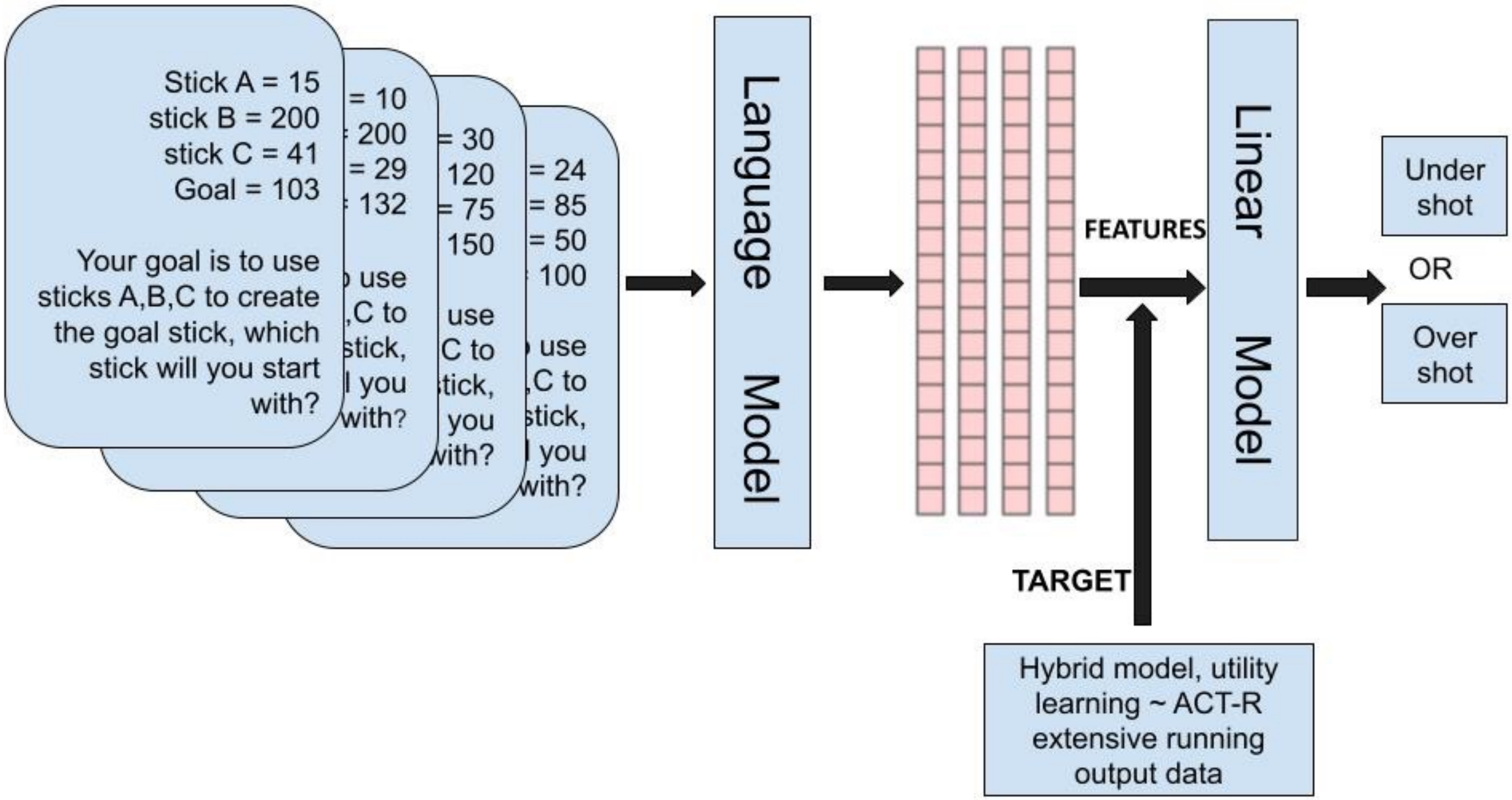
Objective: Integrate ACT-R's neuro-symbolic decision-making with language models to boost performance in domain-specific tasks.

Methodology:

- ACT-R Integration: Employ ACT-R to structure domain knowledge symbolically (declarative chucks and procedural knowledge) for decision-making processes.
- Language Model Fine-Tuning: Adapt a language model (e.g., LLaMA) for domain-specific prompts, focusing on final layer embedding extraction and fine-tuning using ACT-R data from 10,000 runs.
- Data Generation with ACT-R: Generate a dataset of ACT-R decision patterns through extensive runs (10,000 iterations) on problem-solving tasks.
- Model Training: Train a logistic regression model with Lasso 1 regularization, correlating them with ACT-R outputs.

Expected Outcome: This project aspires to advance language models into more cognitively attuned systems, thereby enhancing the capabilities of neural-symbolic AI models.

Illustration of Methodology



Deployment and Generalization

Industry/military training case

Training a chatbot embedded with neural symbolic AI decision making intelligence
Step 1: build the ACT-R model that emulate real time conduction (speed and height control with aviation)
Step 2: prompt the LM with conduction-oriented question like “how to start a helicopter”
Step 3: extract the LM’s last layer of embeddings of the output of the generation case.
Step 4: extensively running ACT-R “ how to start a helicopter” model, collect ACT-R out put.
Step 5: training the linear regression model using embeddings and corresponding ACT-R choices. These embeddings were obtained by passing prompts through the LLaMA model and extracting the hidden activations of the final layer.
Step 6: Use the trained model to complete the Q&A in skills acquisition training.

Bhattacharyya, S., Davis, J., Vogl, T., Fix, M., McLean, A., Matessa, M., & Smith-Velazquezent, L. (2015). Enhancing Autonomy with Trusted Cognitive Modeling. *Association for Unmanned Vehicle Systems International*.

Footnote: Inspired by the aviation training case from Dr. Siddhartha Bhattacharyya.

Conclusion

My work says my heart

My passion aligns with CBI’s vision

Thank you!