Modeling Human Driving Behavior Using ACT-R

Raj Kaneriya, Julian Gay, Ben Barnett, Dylan Burkholder

The Pennsylvania State University

CMPSC 497: AI, Humans and Society

Dr. C. Dancy

May 2 2025

Abstract

This paper provides a simulation-based analysis of driver behavior simulated through human-like actions using ACT-R cognitive modeling. Our goal is to evaluate the effect of cognitive load on drivers' decision-making in dynamic scenarios and to simulate these behaviors through rule-based AI agents. The model also integrates environmental factors and internal states to evaluate how people respond in stressful situations. Additionally, we discuss the ethical and sociopolitical consequences of cognitive modeling and how it relates to the Standard Model of the Mind and sociogenic accounts of human activity.

1. Introduction

The task of driving involves significant cognitive complexity, requiring rapid perception, decision-making, and motor responses under constantly changing conditions. This project explores the use of the Adaptive Control of Thought, Rational (ACT-R) cognitive architecture to model human driving behavior, especially under cognitive stress conditions. The main objective is to develop a model that is capable of simulating human-like responses to unexpected driving situations and changes in the environment, in addition to addressing the wider socio technical and ethical implications involved in developing cognition-centered systems.

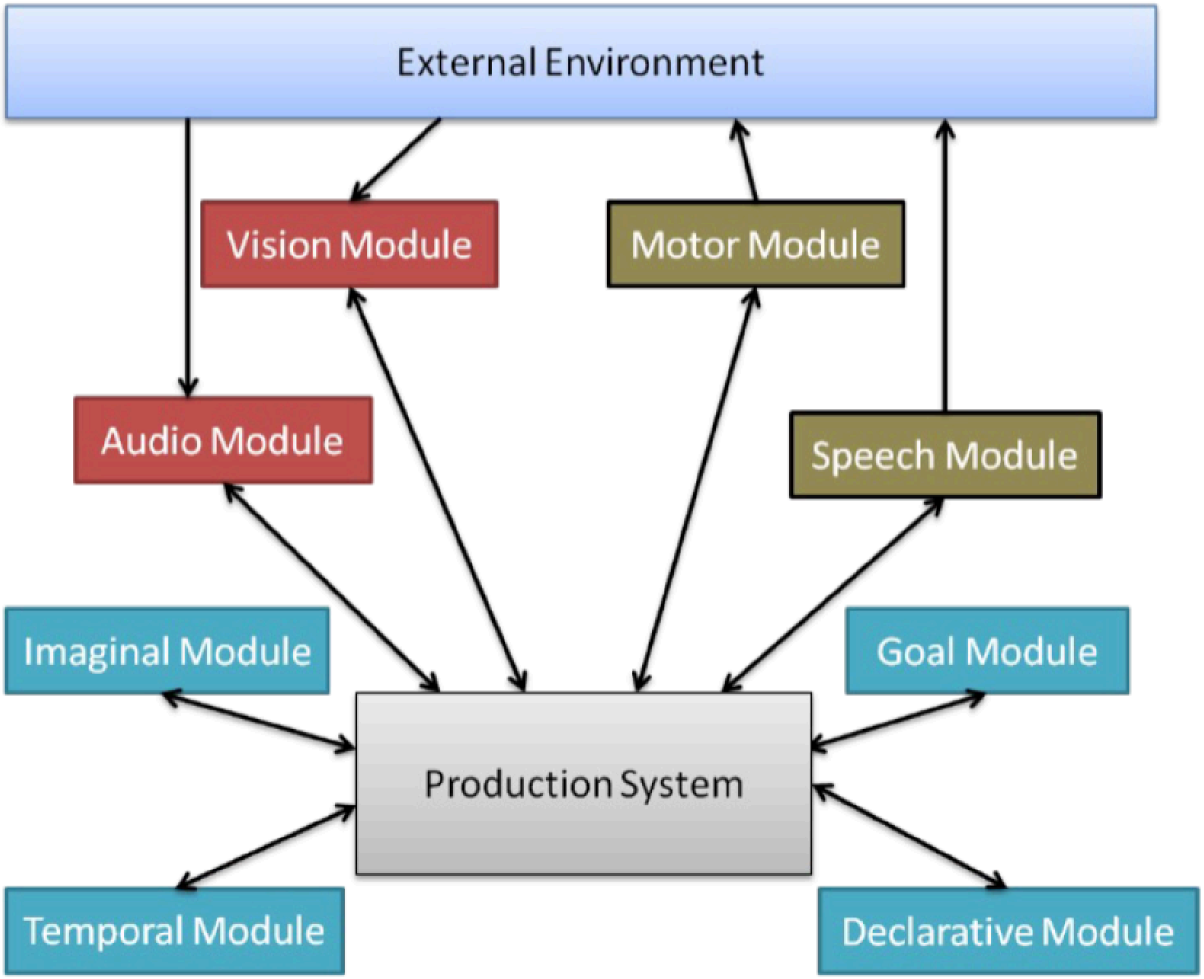
ACT-R, a model devised by John R. Anderson and colleagues, is both a symbolic and a sub-symbolic model based on emulating human memory retrieval, rule following, and task performance through a parallel and modular-processing model (Anderson, 2007). By embedding this framework within a Python-based driving simulation, we investigate how a human-like agent adapts its responses to obstacles, distractions, and varying road conditions, providing insight into cognition, safety, and the construction of inclusive AI.

2. Project Goals

This project is motivated by a set of basic aims. The main aim is to simulate human-like decision-making through the ACT-R model to measure to what degree artificial agents can mimic human rationality in driving situations. A second goal is to study the effect of declarative memory and production rules in controlling behavior in driving scenarios. Another important aim is to investigate how different levels of cognitive load, such as distractions and multitasking conditions, influence driver performance. Additionally, we aim to gain insight into how behavioral tendencies vary in relation to environmental complexity levels. Finally, a critical aim of this work is to examine the ethical implications that arise when human behavior is modeled and replicated in artificial intelligence systems.

3. Description of the System

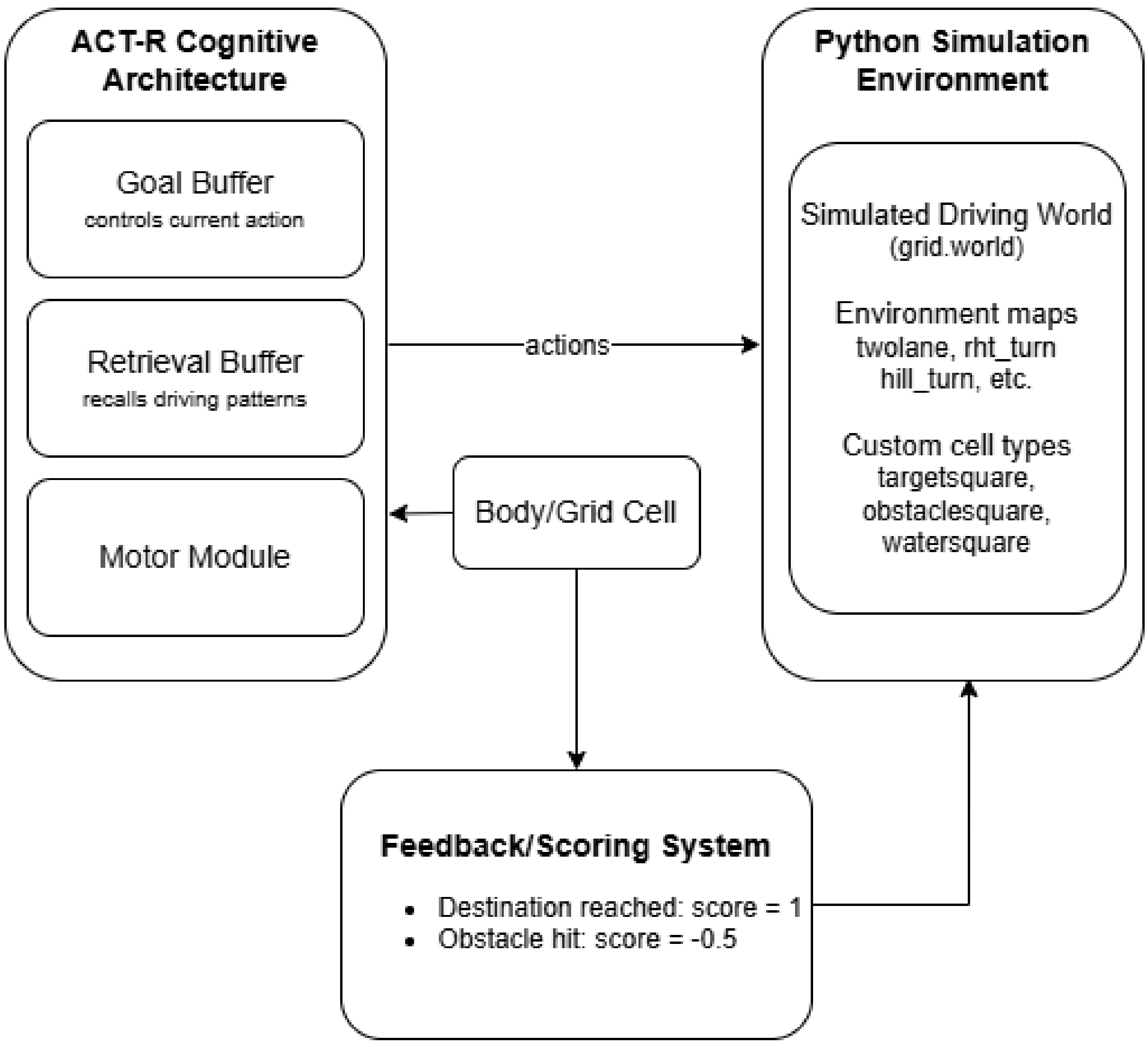
The simulation environment is an abstracted model of a driving environment, written in Python. Within this framework, the ACT-R agent interacts with virtual stimuli in a manner analogous to human cognition. It draws on declarative memory to access past experience, such as pedestrian crossings and unexpected hazards. The agent will apply production rules to work out an appropriate response in such situations, such as slowing down or performing evasive actions. The agent also constantly looks to evaluate both its external environment and internal state, and in response to changes in cognitive load, will adapt its strategy. These processes are facilitated through cooperation between different modules responsible for perception, memory access, motor implementation, and goal management. By manipulating internal parameters like attention demands and fatigue levels, we are able to simulate a range of human cognitive states.



4. Environment

The simulated environment involves the underlying elements typical of a driving scenario. These elements include road networks and lanes of traffic that are navigated by the agent, along with pedestrians, other cars, and unseen hazards that contribute variability and difficulty. Additionally, the model has variable weather conditions that affect visibility and overall safety. Additionally, in-vehicle distractions such as the simulated use of a cell phone add a further complexity. By controlling these variables, the simulation allows an examination of how increasing complexity affects ACT-R agent behaviour and performance.

To illustrate the interaction between the cognitive architecture and the simulated environment, the flowchart below provides a high-level overview of our system design. The ACT-R model consists of components such as the Goal Buffer, Retrieval Buffer, and Motor Module, which collectively drive decision-making and behavior execution. These actions are then processed by a grid-based simulation environment composed of customizable cell types (e.g., obstacles, targets). Feedback is relayed through a scoring system, allowing the agent to adapt based on success or failure outcomes. This design emphasizes the modular relationship between perception, action, and learning in a dynamic world.



5. Adaptation

The ACT-R agent in this simulation demonstrates adaptive behavior in a pair of ways. The agent learns from memory in the sense that it remembers past interactions and uses them to inform later responses under analogous circumstances. It is also cognitively load-sensitive. As the cognitive workload rises from distraction or environmental complexity, the agent changes its behavior to use more conservative and safe strategies, replicating the changes that take place in human behavior under stress (Salvucci, 2006). Third, the agent is sensitive to changes in its world; as traffic density, weather, or visibility evolve, its decision strategy is revised in real time. Finally, the model assesses risk as a function of its internal state. For example, decreased attention or heightened fatigue results in altered decision criteria, influencing the way risks are traded off and the choices made.

6. Ethical and Social Considerations 6.1 Problematic Social Systems and Practices
Modelling human cognition inevitably creates ethical problems, especially when models are used to guide decision-making in artificial systems. One major issue is the potential for moral dilemmas in automated reactions, e.g., choices that must weigh passenger safety against pedestrian safety. In addition, biases built into the cognitive architecture, such as those promoting a rational, Western model of decision-making, risk excluding many cultural or emotional behaviors (Wynter, 2003). Another concern is overreliance on simulation data. Models trained exclusively in simulated worlds may lack the inherent unpredictability and richness of real-world driving. In response to the above concerns, our simulation strives to merge heterogeneous behavioral data and critically engages with the assumptions of the model. Specifically, we reflect on how the "Man genre," a rationalist and exclusionary concept of human identity, informs design choices in AI. This awareness informs a more socially responsible and inclusive modeling process (Cave & Dihal, 2020).

7. Theoretical Connections

7.1 Technological Politics

Langdon Winner argues that all technologies encode political values, and this principle is central to our project. The ACT-R simulation reflects assumptions about what constitutes rational, safe, and autonomous behavior. These assumptions have real consequences, particularly when encoded into AI systems that interact with the public. The question we ask is whether AI should reproduce human cognitive flaws, or if it should be optimized to exceed human limitations.

7.2 Cognitive Science Module

John Anderson's "How Can the Human Mind Occur in the Physical Universe" (2007) outlines human cognition's modular architecture. This is the theory that individual modules handle perception, memory, motor control, and goal processing. Our simulation is founded on this architecture, as the ACT-R agent handles driving tasks due to integrated but separable cognitive functions. Driving requires the simultaneous activation of multiple cognitive resources, so it is a suitable domain for testing modular architectures like ACT-R.

7.3 AI and the Human

Sylvia Wynter's sociogenic code theory suggests human activity is shaped not only by drives that are biologically derived but also by cultural stories. This perspective explains how determinants of human behavior are governed through a struggle between rational thinking and instinct and culturally derived norms and laws and social demands. We add such components to our ACT-R model to improve realism in our simulations of human response. By incorporating behavioral norms such as road etiquette and compliance with legislative norms, we demonstrate culture's influence on cognition.

7.4 AI in Contexts

In "The Whiteness of AI," Cave and Dihal (2020) discuss artificial intelligence's tendency to mirror restrictive cultural norms. Without stringent analysis of such norms, AI systems have the potential to overlook or superficially address cultural diversity. Our work in this project is aimed at addressing such a criticism by designing an ACT-R agent contradicting traditional assumptions about cognition. Our goal is to build a system that is not only technologically competent but also pluralistic in its description of human variability.

8. Standard Model of the Mind

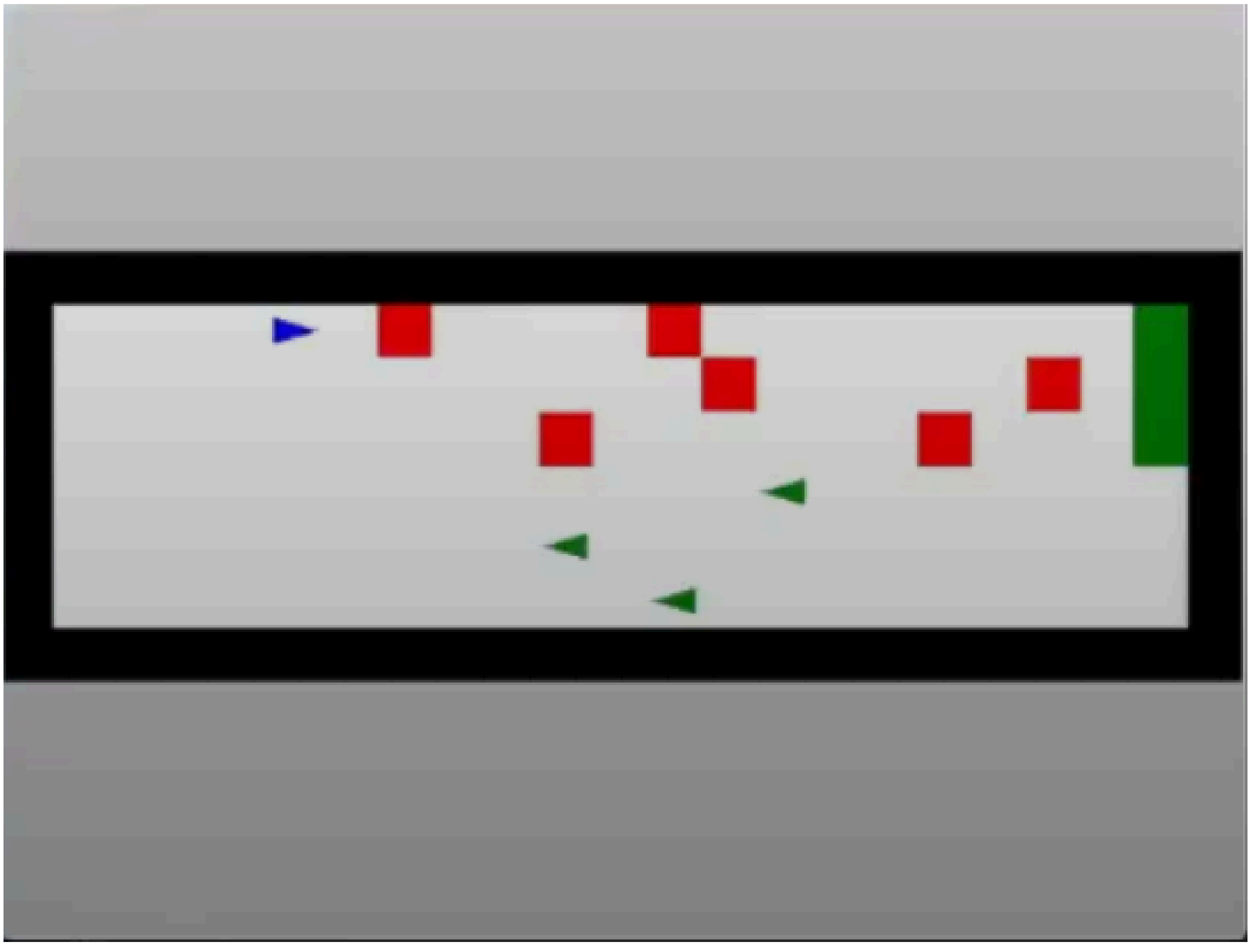
Our simulation is consistent with the Standard Model of the Mind proposed by Laird et al. (2017), identifying key elements of intelligent behavior in perception, cognition, action, and learning. Our perception is achieved through what can be observed and interpreted by the agent via visual inputs. Cognition is derived from interaction between declarative memory and production rules. Actions are carried out through simulated motor functions like braking or turning. Learning is ultimately achieved through modification of declarative memory based on new experiences and influencing future decision-making processes. This compatibility provides strong theory support for our use of ACT-R in driving behavior simulation.

9. The Man Genre

The "Man genre" describes a cultural framework in which human nature is defined in terms of rationality, autonomy, and affective disengagement. This framework has profoundly influenced areas in cognitive science and artificial intelligence studies, sometimes to the exclusion of a more complex or sophisticated view of human activity. Our simulation conducts a critical examination of this genre through its simulation of a wider range of human activities including stress, distraction, and culturally determined behavioral adaptations. This approach supports the development of AI systems that are not only more accurate but also more balanced and socially aware.

10. Evaluation and Results

The development of the simulation revealed several technical and conceptual challenges. Initially, the agent failed to stop appropriately in response to obstacles, requiring adjustments to the production rules and event handling mechanisms. Furthermore, the simulation lacked realistic post-collision behavior. For example, after a crash, the vehicle would continue driving instead of responding appropriately. Another limitation was the difficulty of extracting structured data from the Python tkinter environment, which limited the scope of quantitative analysis. Despite these issues, the system effectively demonstrated the ability to simulate rule-based decision-making under varying cognitive loads. These results support the hypothesis that ACT-R can model dynamic driving behavior in a cognitively plausible way. The image below depicts a running instance of our ACT-R driving simulation. The blue arrow represents the primary agent, whose task is to navigate safely through the environment by avoiding collisions with red square obstacles while remaining within the designated traffic lanes. The green area on the far right denotes the endpoint or goal of the simulation. Green arrows simulate simplified vehicles traveling in the opposite direction, serving as abstract representations of oncoming traffic to introduce additional environmental complexity.



11. Limitations and Future Work

While the simulation provides valuable insights, it is limited by its simplified environment, which does not fully capture the complexity of real-world driving. Agent behavior remains largely deterministic, and future iterations could benefit from incorporating probabilistic models or reinforcement learning techniques to introduce variability. Additionally, the simulation currently draws on a limited dataset, which may not reflect the full range of human driving behavior. Expanding the model to include more diverse behavioral patterns would enhance its representativeness.

Beyond software-based improvements, a promising direction for future work involves connecting cognitive simulations like ours with real-world autonomous vehicle platforms. As companies like Waymo continue developing sensor-rich, AI-driven vehicles, simulations rooted in human cognitive architecture could serve as a complementary layer—offering insight into how humans respond to ambiguous or ethically challenging road scenarios. Incorporating ACT-R-like agents into autonomous system training could help vehicles better anticipate, mirror, or even diverge from human decision-making when appropriate. This integration may be especially beneficial in improving safety, trust, and interpretability in autonomous systems operating in complex social environments.

References

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.

Cave, S., & Dihal, K. (2020). The whiteness of AI. *Philosophy & Technology, 33*(4), 685–703. https://doi.org/10.1007/s13347-020-00415-6

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *AI Magazine, 38*(4), 13–26. https://doi.org/10.1609/aimag.v38i4.2744

Salvucci, D. D. (2006). Modeling driver behavior in a cognitive architecture. *Human Factors, 48*(2), 362–380. https://doi.org/10.1518/001872006777724417

Waymo LLC. (2025). *Waymo Driver*. https://waymo.com/waymo-driver/

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136.

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, after Man, its overrepresentation—An argument. *The New Centennial Review, 3*(3), 257–337. https://doi.org/10.1353/ncr.2004.0015