Toward a More Unified ACT-R Cognitive Architecture

Serhii Serdiuk[1,2][0000-0001-7987-6766] and Frank E. Ritter[1]

[1] College of IST, Penn State, University Park, PA 16802, USA
[2] National University "Zaporizhzhia Polytechnic", Zhukovskoho 64, Zaporizhzhia, Ukraine

Abstract. In this paper we consider what remains to create a unified theory of cognition, updating Newell's [1] report. We examine ACT-R as a new exemplar. After quickly summarizing what has been covered, we create a new schematic for ACT-R to add capabilities and mechanisms for a 5 to 10-year research plan. This includes more meso-level features of cognition and high-level aspects of cognition (that will inevitably increase the functionality while decreasing the usability of ACT-R unless aspects can be turned off). We suggest integrating more aspects of physiology and psychology as well as more social aspects. To accomplish this, we introduce six new functional modules. These additions enable the architecture to account for physiological and emotional conditions, stable personal traits, value-based evaluations, experiential learning, and gradual behavioral change. As a result, this extended ACT-R architecture offers a more comprehensive and flexible model of human cognition, bridging previously underrepresented areas of psychology, physiology, and behavior.

Keywords: ACT-R, unified theories of cognition, cognitive architecture, common model of cognition, P-adic calculus, Mental space.

1 Introduction

ACT-R is the current broadest, best, and most tested architecture for human cognition [2-4]. One of the latest versions of ACT-R [3] captures the essential cognitive mechanisms involved in perception, memory retrieval, goal setting, action selection, and motor execution.

The current system leaves out a lot of psychology, physiology, kinesiology, and biology related to behavior. Some of these areas have been hard to study, but it would be worthwhile to show how they can be integrated in a more serious way than Newell [1] or Anderson [5] provides. This paper shows how many more areas can be represented as extensions to ACT-R, and provides a blue print for the next more complete version of ACT-R as an exemplar UTC.

We draw on our review of perception [6], Kotseruba and Tsotsos' reviews of architectures [7,8], the common model of cognition [9], and discussions with a range of psychologists and other -ologists. We start by reviewing the extensions grouped by major area. Fig. 1 shows the revised architecture.

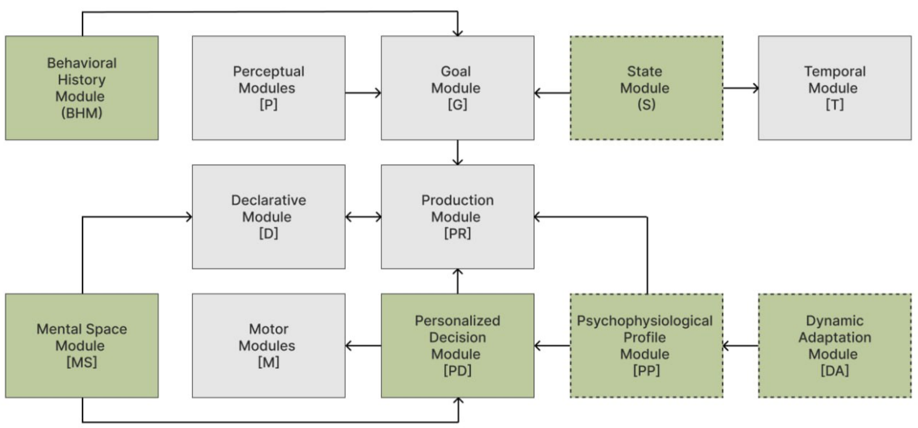

Fig. 1. Expended ACT-R schematic. The new modules have green fill. Newly proposed modules not functionally engaged in the example scenario have dashed borders.

Several modules are not illustrated because the example task—a comparison of two bus drivers performing typical driving tasks—focuses on a normal, non-extreme driving situation, where the agents operate under stable physiological and psychological conditions. As a result, modules such as the PP, S, and DA remain inactive or exert only baseline influence, and their dynamic mechanisms are not demonstrated in detail. These modules are nonetheless included in the extended architecture to represent the agent's capacity for modeling stress, fatigue, long-term adaptation, and individual differences, which may be important for simulating situations with heavy workload, sudden events, or long periods of task performance.

While the current version of ACT-R [3] incorporates essential modules for perception, memory retrieval, goal setting, production selection, and motor execution, certain functional areas—such as internal cognitive states, long-term adaptation, and behavioral history—remain beyond the scope of this architecture. Specifically, the architecture does not include detailed modeling of low-level physiological processes, such as hormonal, metabolic, or cardiovascular mechanisms, which would require integration with complex physiological simulators like HumMod [10].

Furthermore, the classical ACT-R architecture only models vision, hearing, and manual actions, and does not include other sensory systems [7]. As noted by Ritter and Serdiuk [6], a relatively complete list of human senses includes up to 24 distinct sensory modalities, encompassing both external (e.g., touch, temperature) and internal (e.g., hunger, fatigue, emotions) senses—all of which remain beyond the scope of the classical ACT-R model.

These limitations highlight the need to extend the classic ACT-R framework to better account for these factors. To address these gaps, we propose an expanded architecture that introduces several new modules specifically designed to capture these aspects of cognition and behavior.

Figure 1 shows how the proposed modules can integrate into the current version of the ACT-R architecture [3]. Table 1 provides a brief description of these new modules.

Table 1. New modules in the extended ACT-R architecture.

Module	Function	Notes / Interactions
Psychophysiological Profile	Stores stable individual characteristics: reaction time, stress tolerance, impulsivity, conformity.	Influences P, G, Matching/Selection (Production Cortex), PD, and M modules; gradually updated by DA module.
Mental Space	Maintains a personal p-adic structure of meanings and associations; categorizes ideas as "interesting," "forbidden," or "doubtful."	Modulates D, Retrieval, G, PD, and Production (Matching/Selection) modules; updated by BHM
Personalized Decision	Integrates inputs from S, PP, MS, and BHM modules to select the next action; directly modulates PR to bias action selection according to personalized criteria.	Receives data from S, PP, MS, and BHM modules; indirectly influenced by DA module via updates to the PP; modulates Production (Matching/Selection).
Dynamic Adaptation	Gradually adjusts stable traits in the PP module based on long-term patterns.	Receives data from S and BHM modules; influences all modules via updated profile parameters.
State	Represents the agent's temporary internal state: stress, fatigue, emotional arousal, alertness.	Affects P, G, Production (Matching/Selection), and T modules; informs DA module.
Behavioral History	Records past actions and outcomes, building a behavioral experience base.	Biases Production (Matching/Selection) and G modules; updates MS; informs DA module.

2 P-Adic Representation of Mental Space

For the formal representation of information at the human sensory inputs and its cognitive processing, we will use p-adic hierarchical trees, proposed by Khrennikov [11,12], as a mathematical model of a mental space. According to Khrennikov [12], thinking systems operate with I-sequences, which the author called I-states. The set of I-states, X_I, has the structure of an m-adic tree and forms the mental space Z_m: $X_I = Z_m$. Hierarchical families of I-states form I-objects, which are proposed to be called "associations" (bundles of branches with a common root). Families of associations form higher-level I-objects, which are called "ideas."

Two I-states, x and y, in the mental space will be closer the longer their common root is. This topology is described by the metric $\boldsymbol{\rho_m}$. Khrennikov [12] represents the mental space as an ultrametric space.

Let $m > 1$ be a fixed natural number. Then, the ultrametric can be represented as [12]:

$$\boldsymbol{\rho_m}(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}|_{\boldsymbol{m}}, \tag{1}$$

where $|.|_m$ is the so-called m-adic norm (analogous to the modulus for real numbers).

The norm is defined as follows. Let $\boldsymbol{x} = (\boldsymbol{a_0}, \boldsymbol{a_1}, \dots, \boldsymbol{a_n} \dots)$. The point x = $(\boldsymbol{a_0}, \boldsymbol{a_1}, \boldsymbol{a_2}, \dots, \boldsymbol{a_n}, \dots)$ in the space Z_m can be identified with a natural number:

$$\boldsymbol{x = a_0 a_1 \dots a_k \dots \equiv a_0 + a_1 m + \dots + a_k m^k} \tag{2}$$

Then, $\boldsymbol{\rho_m(x, y)} = \frac{\boldsymbol{1}}{\boldsymbol{m^k}}$, if $\boldsymbol{a_j = 0, j = 0, \dots, k-1,}$ $\boldsymbol{and\ a_k \neq 0.}$ In other words, to find the distance between x and y, it is necessary to find the first position k where x and y have different digits.

In addition, Khrennikov's architecture includes databases of interesting, forbidden, and doubtful ideas, reflecting the experience of a particular person [12]. We will use them in the MS module.

3 Simulation of Agent Decision-Making in Extended ACT-R

Let us examine the operation of several aspects of the extended ACT-R architecture using the example of two agents—bus drivers. To assess the psychological characteristics of drivers, we can use a battery of well-known psychological tests, theories and tools.

For example, as a result of testing, the following characteristics were obtained. Driver 1 (Ivan) demonstrates an average reaction time of 250 ms, high stress tolerance, high risk propensity, low conformity (not inclined to follow rules), and medium impulsivity. He enjoys driving fast, and violations do not cause him internal stress. Driver 2 (Alex) shows an average reaction time of 350 ms, medium stress tolerance, low risk propensity, high conformity, and low impulsivity. He prefers to follow rules and enjoys calm and safe driving. (Behavioral histories for both drivers are presented later in Table 7.)

Table 2 provides a fragment of the description of perceived objects in the P module (Visual) using 5-adic numbers.

For example, Ivan and Alex are approaching an intersection where the traffic light is not regulating traffic in the usual mode, but is flashing yellow. At the crosswalk in front of them is a pedestrian who is going to cross the road. At the same time, the traffic at the intersection is quite intense—the space is occupied by about 50-80% of cars, but the traffic is not blocked.

According to Table 2, the P module represents objects as follows: Traffic Light (object 1) – 22403; Pedestrian (object 2) – 22300; Markings (crosswalk) (object 3) – 22000; Cars at the intersection (object 4) – 22200.

The *p*-adic coding in the MS module is presented in Table 3. According to Table 3, the value m=5 (as the largest). An example list of ideas for agents in the MS module is presented in Table 4.

Let us consider how the MS module generates interpreted *p*-codes.

Step 1. Identifying situational features. Based on the information about the four input objects, the M module extracts the following (Table 5). Thus, we have final mental *p*-code: 120232.

Table 2. P-adic code table for perceived objects.

Code	Positions in p-adic number				
	0	**1**	**2**	**3**	**4**
	Distance of perception	**Object brightness/visibility**	**Object type**	**Object subtype (e.g., for signs and traffic light)**	**Specific category (e.g. for regulatory)**
0	very close	very low	road markings	transverse markings	Crosswalk
1	close	low	sign	0 - regulatory	0 - STOP
					1 - YIELD
					2 - Do Not Enter
					3 - No Parking
					4 - SPEED LIMIT
				1 - warning	
				2 - guide	
				3 - information	
2	medium	normal	car		
3	far	high	pedestrian		
4			traffic light	0 – for car	0 - red
					1 - steady yellow
					2 - green
					3 - flashing yellow
					4 – absent/off
				1 – for pedestrian	

Table 3. P-adic coding in the Mental Space module.

Position	Purpose	Examples of values
0	Motivation / intention	0 = risk, 1 = caution, 2 = social responsibility
1	Action strategy	0 = aggressive, 1 = passive, 2 = optimized
2	Situation context	0 = intersection, 1 = highway, 2 = parking
3	Obstacle / road user	0 = none, 1 = car on side, 2 = pedestrian, 3 = child
4	Type of signal / traffic light	0 = red, 1 = steady yellow, 2 = green, 3 = flashing yellow 4 = absent/off
5	Space occupancy	0 = empty, 1 = <50%, 2 = 50–80%, 3 = busy

Table 4. List of ideas for agents.

Category of idea	P- code	Explanation
Ivan (likes to take risks)		
Interesting	000030	Drive through a flashing yellow light, without obstacles, when the road is empty
	020031	Quickly assess the situation and take a risk on a flashing light, with minimal traffic
	001142	Aggressively overtake on the highway with moderate traffic, despite a car on the side and a traffic light that is off
Doubtful	120012	Slow down, but if possible, slip through on a yellow light, with moderate traffic
	010031	Slow down a little out of politeness on a flashing light, with few cars
Forbidden	100000	Coming to a complete stop before a red light, even if everything is clear - wastes time
	212223	Give way to a pedestrian in a parking lot when the traffic light is green (for cars) and the space is busy
Alex (careful)		
Interesting	120112	Slow down and yield at an intersection if there is a car at the side and the signal is yellow
	220232	Let pass if there is a pedestrian and the yellow light is flashing
	222042	Park carefully if there are no obstacles and the traffic light is off
Doubtful	110012	Try to pass on yellow with caution in moderate congestion
	211142	Overtake slowly on highway with car on side and traffic light off
Forbidden	000030	Passing on a flashing light without stopping, even if it's empty
	212223	Aggressively overtaking on the highway when the light is green, without checking traffic

Table 5. Identified situational features.

Position	Parameter	Output based on objects
0	Motivation	1 (caution) or 2 (social responsibility) — because of the pedestrian and other cars
1	Strategy	2 (optimized) — you need to balance attention: both on the pedestrian and on the cars
2	Context	0 (intersection) — Object 4 points to the intersection
3	Obstacle	2 (pedestrian) — Object 2
4	Traffic light	3 (flashing yellow) - based on Object 1
5	Space occupancy	2 (50–80%) - Based on Object 4: Cars at intersection

Step 2. Comparison of the final mental *p*-code (120232) with the idea databases of Ivan and Alex (Table 4) using 5-adic ultrametrics. Calculations using formula (1) showed that the minimum distance for the current situation is determined to Ivan's doubtful idea 120012 and Alex's interesting idea 120112: $\rho = 0.008$. These ideas will be the result of the MS module's work. All other ideas (except for another doubtful idea

for Alex) differ already in the first position (action type), which gives a greater distance: ρ = 1.

The solution chosen by the MS module may be influenced by the life experience of the agents, represented as records in the BHM module. The structure of records in this module is presented in Table 6.

Table 6. The structure of records in the BHM module.

Situation Code	Action	Result	$r_{i,j}$	$f_{i,j}$
		Ivan		
120012	slow down	safe passage	+0.9	6
120112	slow down	smooth coordination	+0.6	3
100000	stop	unnecessary delay	+0.2	2
120012	stop	short hesitation	+0.3	1
120012	go	pedestrian startled	-0.7	3
001142	go	successful overtake	+0.8	4
		Alex		
120112	slow down	safe passage	+0.8	4
120012	slow down	minor delay	+0.5	2
120112	stop	pedestrian safe	+1.0	3
120012	stop	hesitation, no harm	+0.4	1
120112	go	risky, no incident	-0.3	1

The BHM module operating algorithm is as follows.

Step 1. Filter BHM data by *p*-code: Select only those records where **`situation_code`** matches the p-code from MS. If there is no data at all for a given p-code:

- We look for the closest similar *p*-codes (e.g., differing by one bit).
- Or we use the default value (for example, “stop” in an unfamiliar situation).
- Or we start the fallback mode: the new action is tested and written to the BHM with a low weight.

Step 2. Aggregation of data for each action: For each action a, calculate the aggregate reliability estimate:

$$\boldsymbol{D_a} = \frac{\sum_{i=1}^{n} r_i \times f_i}{\sum_{i=1}^{n} f_i}, \tag{3}$$

where $\boldsymbol{r_i}$ — reward, $\boldsymbol{f_i}$ — frequency.

However, to handle conflict situations, when one action has a high reward but a low frequency, and the other action has the opposite, formula (3) can be modified:

$$\boldsymbol{D_a} = \left(\frac{\sum_{i=1}^{n} r_i \times f_i}{\sum_{i=1}^{n} f_i}\right) \times \boldsymbol{log}\left(\mathbf{1} + \sum_{i=1}^{n} \boldsymbol{f_i}\right), \tag{4}$$

This approach increases the importance of reliable, frequently observed actions, while little-tested but highly rated actions will be inferior to consistently good ones.

Step 3. Once the MS module selects the most cognitively proximal idea, the BHM module is queried to evaluate the reliability of potential actions associated with this mental representation. Selecting the action with maximum D_a: Select $\boldsymbol{a}^* = \boldsymbol{arg\ max_a D_a}$. If $\boldsymbol{D_a} < \theta$, (where θ is a threshold coefficient, for example, $\theta = 0.5$), then signal uncertainty in the action (possibly contact with G module or clarify the strategy with the MS). An example of Ivan and Alex's experience records is presented in Table 7.

The results of calculating the reliability of actions D_a according to formula (4) are presented in Table 8.

Table 7. Example of BHM's records for agents.

Component	Output based on objects
`situation_code`	*p*-adic number (as in MS module) describing the situation
`action_code`	action code (e.g. speed up, slow down, stop)
`result`	outcome of an action (safely / accident / delay / success, etc.)
`reward or feedback`	numerical evaluation of the result of an action (e.g. from -1 to +1)
`frequency`	the number of times this behavior was observed in a similar situation

Table 8. The results of calculating the reliability of actions.

Situation Code	Action	$r_{i,j}$	$f_{i,j}$	Contribution	D_a
		Ivan			
120012	slow down	+0.9	6	5.4	1.842
120112	slow down	+0.6	3	1.8	
100000	stop	+0.2	2	0.4	0.323
120012	stop	+0.3	1	0,3	
120012	go	-0.7	3	-2.1	0.327
001142	go	+0.8	4	+3.2	
			Alex		
120112	slow down	+0.8	4	3.2	1.362
120012	slow down	+0.5	2	1.0	
120112	stop	+1.0	3	3.0	1.368
120012	stop	+0.4	1	0.4	
120112	go	-0.3	1	-0.3	-0.208

According to the data in Table 8, the action "slow down" will be chosen for Ivan, supported by MS module and reinforced by BHM confidence (D_a = 1.842). Alex hesitates between "stop" and "slow down": both options are well supported. However, "stop" has a slightly higher score (D_a = 1.368) and can be chosen as a safer option. As the reliability values exceed the internal decision threshold θ, this causes the activation

of the corresponding procedural rules in the PR module. Thus, ideas 120012 (Ivan) and 120112 (Alex) selected by the MS module are confirmed.
This causes the activation of the corresponding rules in PR module. For example, for Ivan (Figure 2):

```
IF situation_code = 120012
AND reliability(D_slow down) > θ
THEN action = slow down
```

Fig. 2. Example of a production rule for situation 120012.

Interpretation of the rule (Figure 2):

- **`situation_code = 120012`**: *p*-adic interpretation of the current situation (derived from the Perceptual modules and processed by the MS module).
- **`reliability(D_slow_down) > θ`**: is the result of the analysis performed by the BHM module, reflecting accumulated experience.
- **`THEN`**: the motor action is executed.

Once the procedural rule has been activated and the action "slow down" is selected, the M module is responsible for translating this decision into executable motor commands.

This representation can be overlaid upon ACT-R's declarative memory system, in the same way that other memory systems such as holistic memories have been created.

4 Conclusion

We can now imagine a much more complete computational model of human behavior—one that incorporates many more aspects than Newell or Anderson originally proposed. The extended ACT-R architecture introduced in this paper demonstrates how additional modules can formalize numerous concepts from psychology that have long been acknowledged but rarely represented in cognitive architectures. These include personal style, personality traits, stress resilience, impulsivity, risk sensitivity, long-term behavioral adaptation, and the subjective interpretation of experience.

By integrating these elements, the model paves the way toward simulating not only general human cognition, but also how different individuals think, feel, adapt, and decide under varying circumstances. This enables a richer representation of behavior across time, context, and task demands. We also propose a way to start to implement some of these concepts.

Importantly, the proposed extensions maintain compatibility with the modular, testable principles of ACT-R, while allowing researchers to incorporate latent psychological variables that previously remained outside its scope. This invites further work on experimental validation, refinement of mental-space encoding, and integration with models of physiology, affect, and personality, moving us closer to a truly unified theory of situated human cognition.

One important methodological limitation of the proposed approach concerns the *p*-adic representation of mental space. In such representations, digits in the higher positions (closer to the root) dominate the calculated distance, making the system overly sensitive to differences in early features while ignoring distinctions in less significant positions. This property may distort the relative importance of perceptual or semantic features depending on their encoding position. To address this limitation within the extended ACT-R framework, the PD module can apply additional weighting or normalization procedures to adjust the influence of features depending on context or task demands, effectively compensating for position dominance in the *p*-adic metric.

Another limitation is that these extended capabilities provide a fuller representation but will also make ACT-R more complex. Usability will have to be kept in mind.

Future work includes extending the architecture to model extreme or stress-inducing scenarios, incorporating broader sensory systems as suggested by Ritter and Serdiuk [6], and developing physiological simulators for more accurate representation of internal agent states.

References

1. Newell, A. Unified theories of cognition. MA: Harvard University Press, Cambridge (1990).
2. Chong, Hui-Qing, Tan, Ah-hwee, and Ng, Gee-Wah. Integrated cognitive architectures: A survey. (2009).
3. Trafton, J. G., Hiatt, L. M., Harrison, M.A., Tamborello, F. P., Khemiani, S. S., & Schultz, A. C. ACT-R/E: an embodied cognitive architecture for human-robot interaction. Journal of Human-Robot Interaction 2 (1), 30–55 (2013).
4. Jiménez, J.P., Martin, L., Dounce, I.A. et al. Methodological aspects for cognitive architectures construction: a study and proposal. Artif Intell Rev 54, 2133–2192 (2021).
5. Anderson, J. R. How Can the Human Mind Occur in the Physical Universe? Oxford University Press, New York (2007).
6. Ritter, E. F., Serdiuk, S. Towards a comprehensive summary of senses for cognitive architectures. In 57th Annual Meeting of the Society for Mathematical Psychology (MathPsych) and the 22nd International Conference on Cognitive Modeling (ICCM), pp.152–158.Tilburg University, the Netherlands (2024).
7. Kotseruba, I., & Tsotsos, J. K. 40 years of cognitive architectures: Core cognitive abilities and practical applications. Artificial Intelligence Review, 53(1), 17–94 (2020).
8. Kotseruba, I., & Tsotsos, J. K. (in press). The computational evolution of cognitive architectures. Oxford University Press, Oxford, UK.
9. Stocco, A., Sibert, C., Steine-Hanson, Z., Koh, N., Laird, J. E., Lebiere, C. J., et al. Analysis of the human connectome data supports the notion of a "Common Model of Cognition" for human and human-like intelligence across domains. NeuroImage, 235, 118035 (2021).
10. Brener, M., Becera, N., Pruett, D., Ritter, F. E., Dancy, C. L., & Webster, I. Manual for HumMod (Salt Version 3.0.4) (Tech. Report No. ACS 2019-1, Version: 2.9). Applied Cognitive Science Lab, College of Information Sciences and Technology, Penn State. (2019).
11. Khrennikov, A. Quantum-like formalism for cognitive measurements, BioSystems 70, 211–233 (2003).
12. Khrennikov, A. Modeling of thinking processes in p-adic coordinate systems [in Russian]. FIZMATLIT, Moscow (2004).