KAIST FALL 2025

CS473: INTRO TO SOCIAL COMPUTING

SOCIAL.CSTLAB.ORG

Lecture 19: Simulations (part 2)

2025.12.04

Joseph Seering

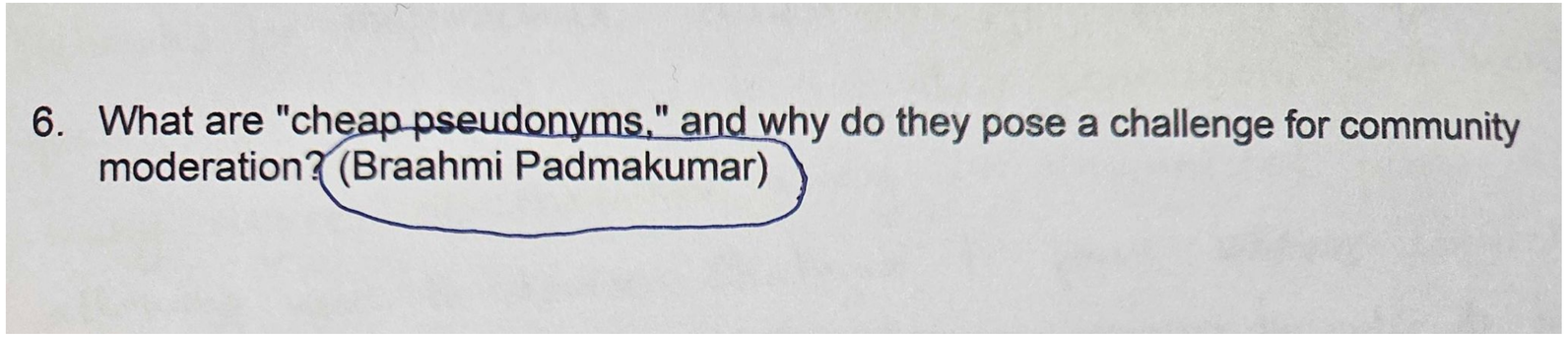
PREVIEW OF DPM5: FINAL PRESENTATIONS

DPM5 studio (Final Presentations) on December 11th in this room!

- 5 minutes to present the problem, the core tasks you're trying to support, how your solution works and how it's better than alternatives, and how your deployment went.
- There is no Q&A during presentations.
- After presentations, there will be a 30 minute demo session where everyone in the class will try what you built.

FINAL EXAM

- Next Tuesday, the exam question bank will be released.
- The exam will be entirely made of (slightly modified) questions you submitted.

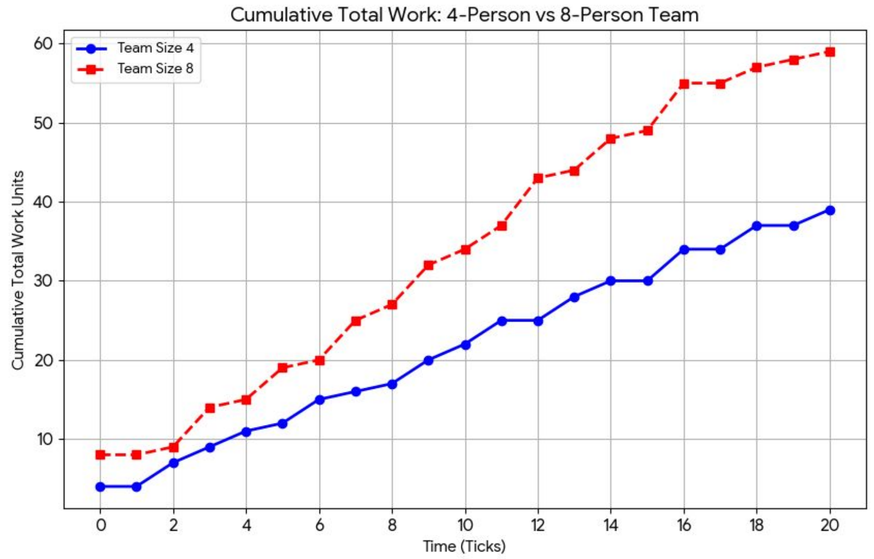

6. What are "cheap pseudonyms," and why do they pose a challenge for community moderation? (Braahmi Padmakumar)

Activity: Make a simple agent-based simulation

https://bit.ly/CS473GroupProjectSimulation

My group work model

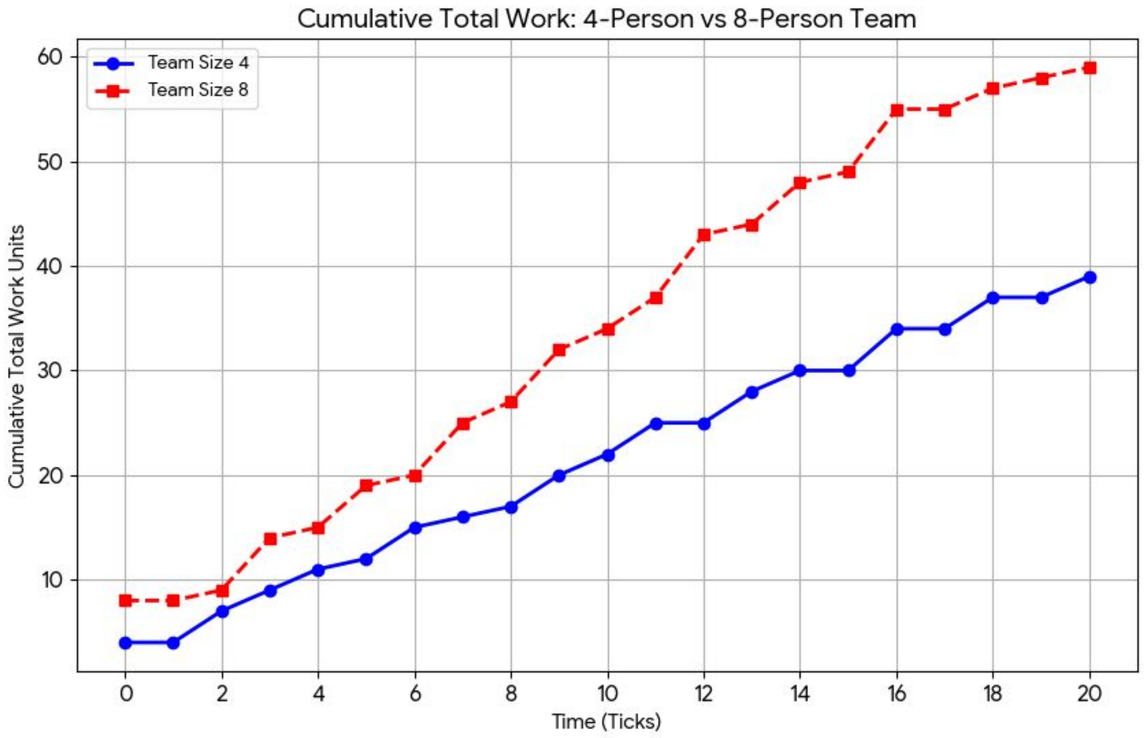
$$P_t = \left(1 - \frac{\text{Agents Working}_{t-1}}{N+1}\right) \times \frac{1}{\sqrt{N/2}}$$



Time	Probability	Agents Working	Ag_1	Ag_2	Ag_3	Ag_4
0	-	**4**	1	1	1	1
1	0.141	**0**	0	0	0	0
2	0.707	**2**	1	0	1	0
3	0.424	**1**	1	0	0	0
4	0.566	**2**	0	0	1	1
5	0.424	**1**	1	0	0	0
6	0.566	**4**	1	1	1	1
7	0.141	**0**	0	0	0	0
8	0.707	**3**	1	1	1	0
9	0.283	**0**	0	0	0	0
10	0.707	**4**	1	1	1	1

Time	Probability	Agents Working	Ag_1	Ag_2	Ag_3	Ag_4	Ag_5	Ag_6	Ag_7	Ag_8
0	-	**8**	1	1	1	1	1	1	1	1
1	0.056	**0**	0	0	0	0	0	0	0	0
2	0.500	**6**	1	1	1	1	1	1	0	0
3	0.167	**1**	0	0	0	0	0	0	0	1
4	0.444	**4**	0	1	0	1	1	0	0	1
5	0.278	**3**	0	0	1	0	0	1	1	0
6	0.333	**1**	0	0	0	0	1	0	0	0
7	0.444	**4**	0	0	1	0	0	1	1	1
8	0.278	**4**	1	0	0	1	1	0	1	0
9	0.278	**0**	0	0	0	0	0	0	0	0
10	0.500	**0**	0	0	0	0	0	0	0	0

Is this accurate?



Why do we care about simulations?

1. Building a simulation helps us learn about the thing we're simulating.
   - The exercise of building a simulation forces us to think really deeply about how something works.

2. Ideally, *prediction.* Can we simulate complex systems in enough detail to make predictions?

3. How is this different from modeling?

So what next?

How do LLMs change agent-based modeling?

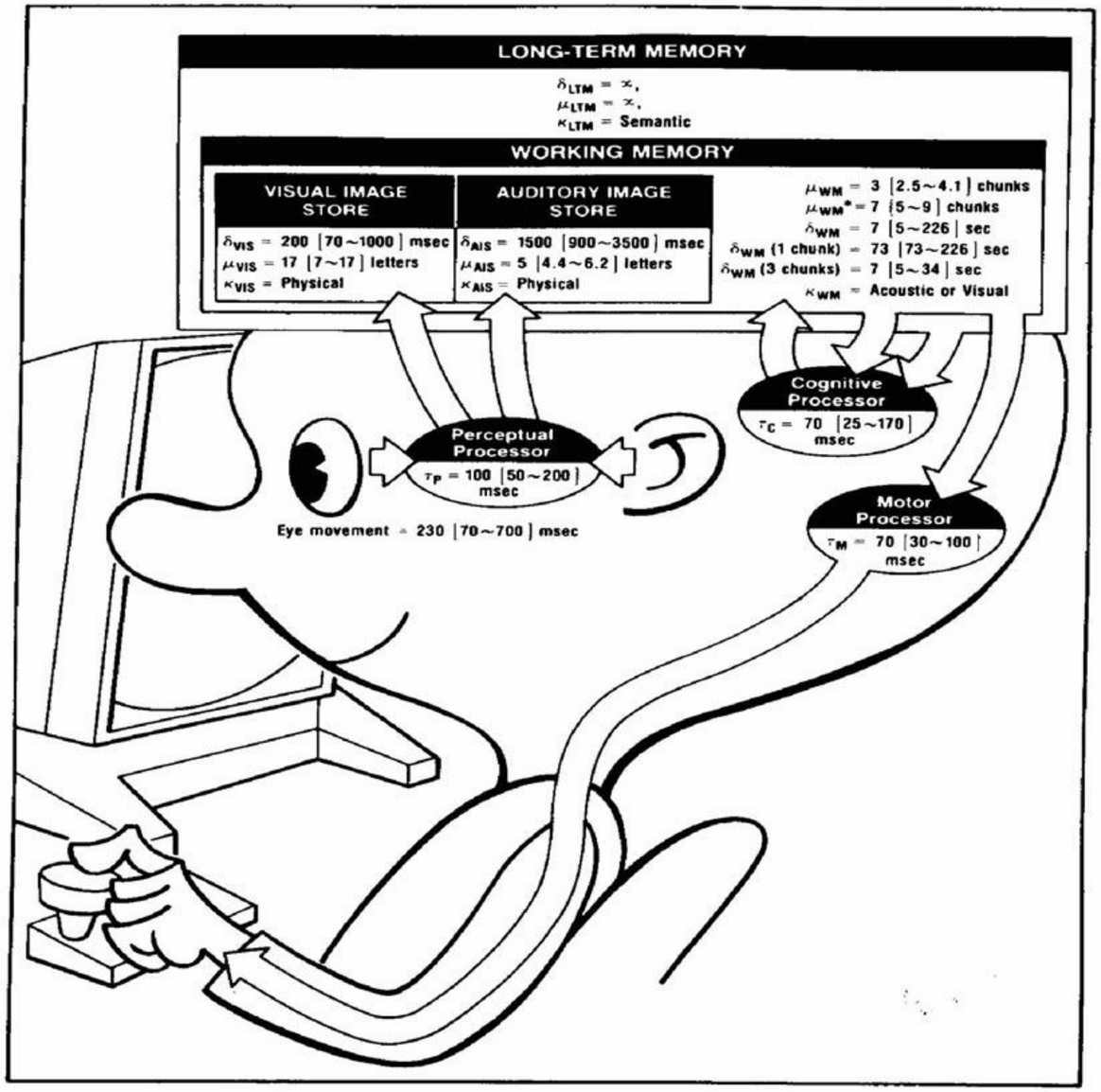
How can we simulate complex populations, now that we aren't limited to simple rules?

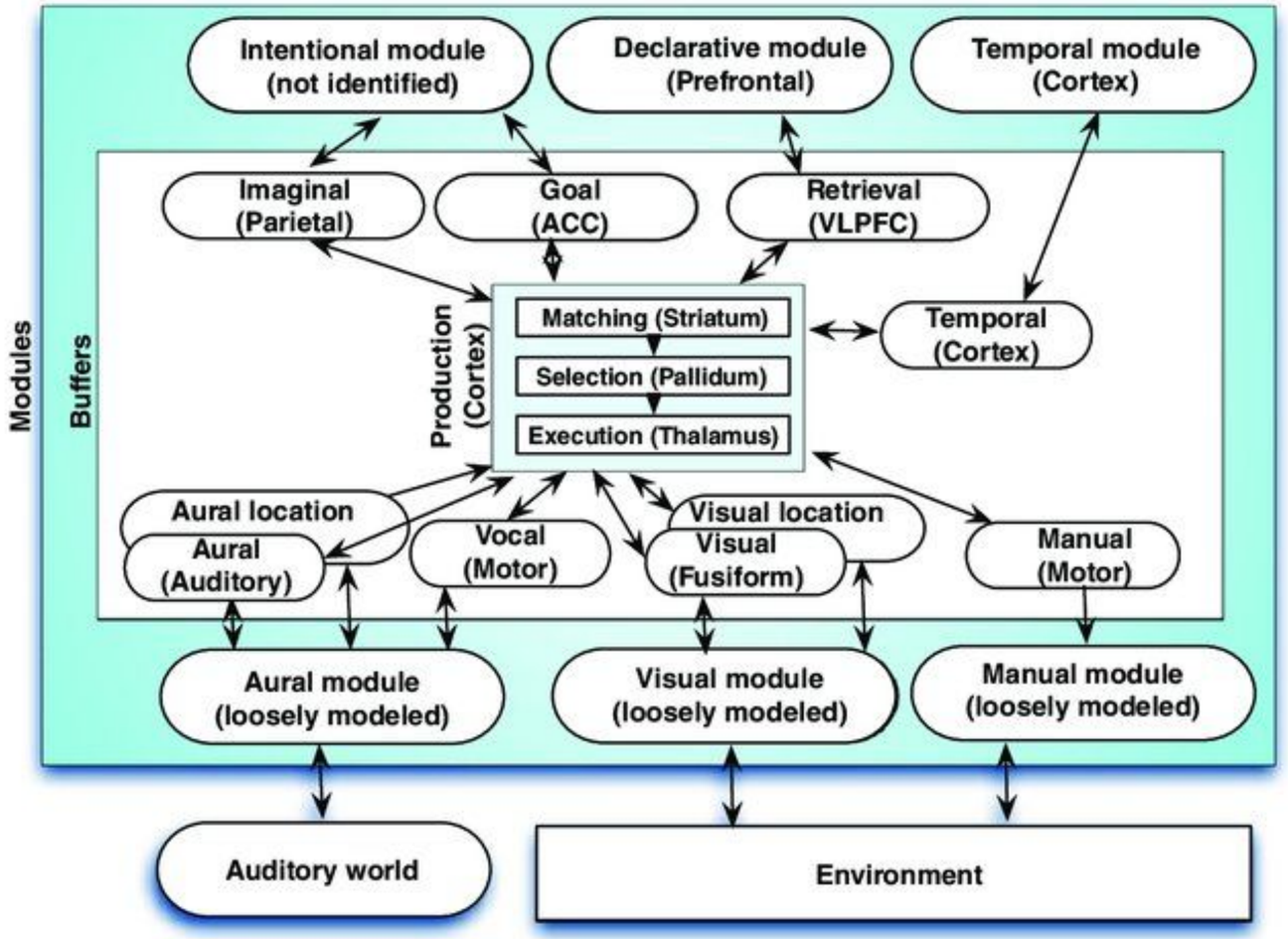
A simple approach would just be… Ask ChatGPT. But can we do a little better?

Cognitive Architectures

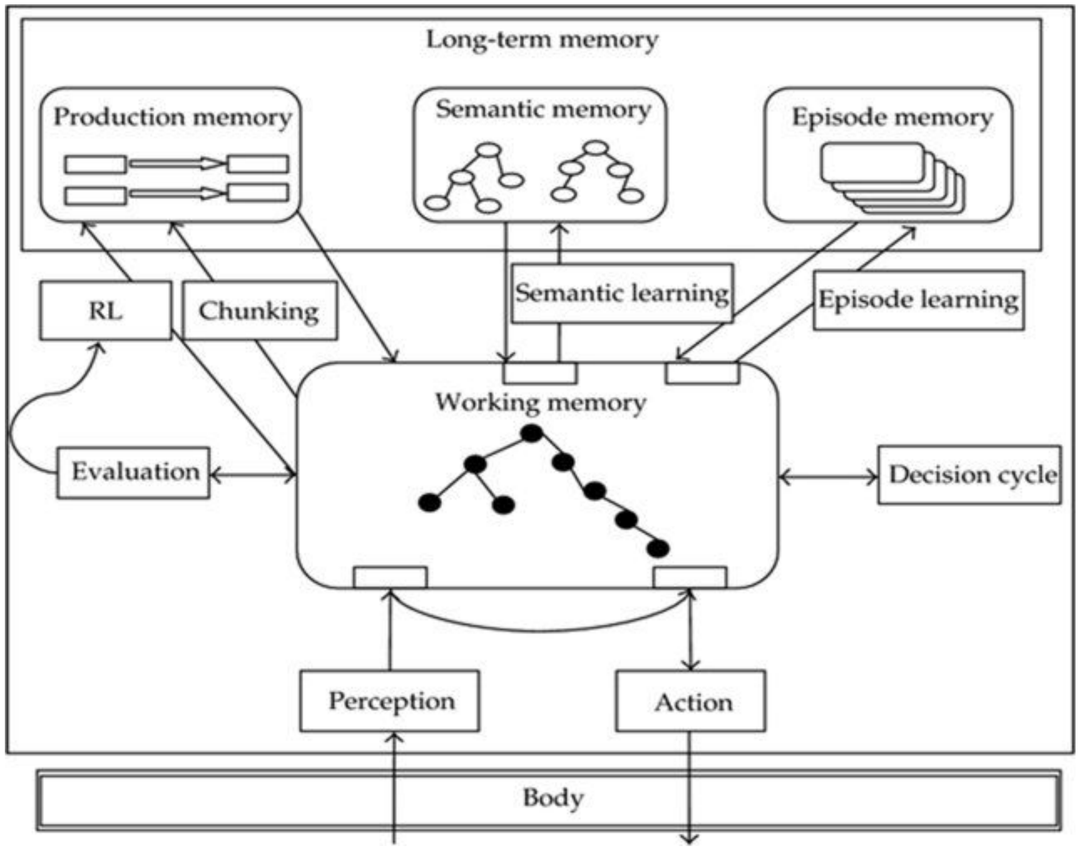
We can replace simple rules with (basically) simulated brains.

So how do we simulate a brain?

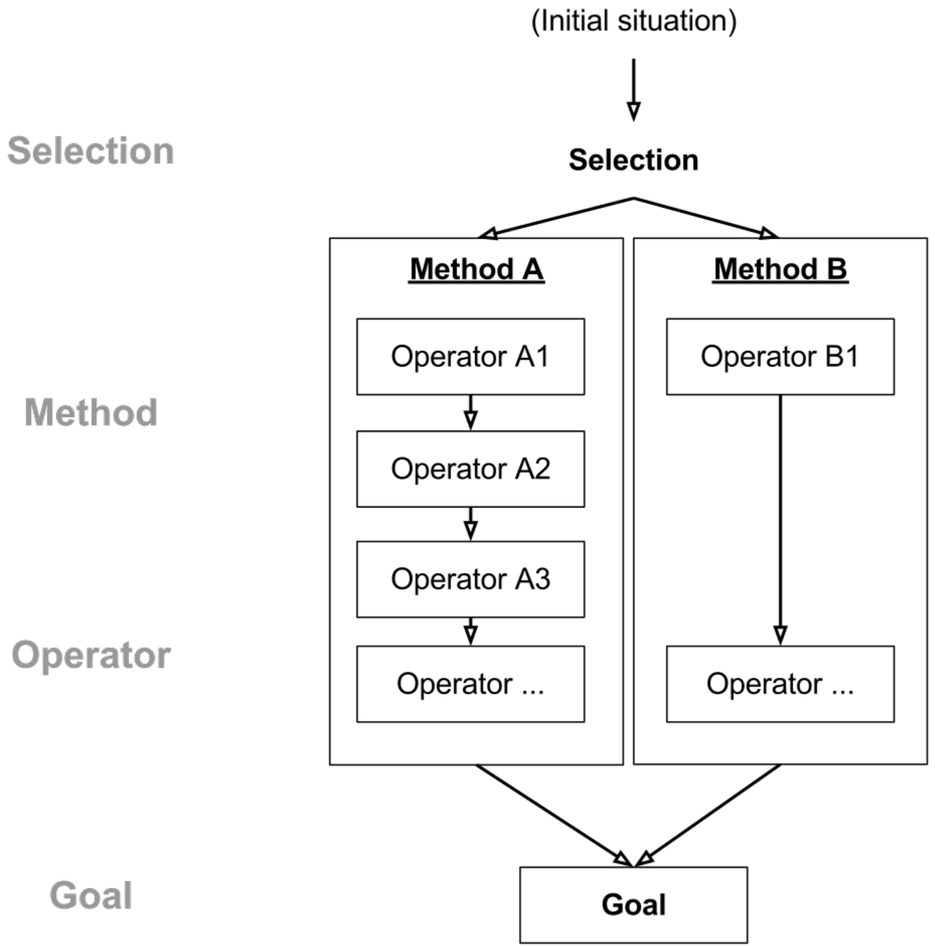




ACT-R model



SOAR model


(Initial situation)
Selection
Selection
Method
Operator
Goal
Method A
Method B
Operator A1
Operator A2
Operator A3
Operator ...
Operator B1
Operator ...
Goal

GOMS model

Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park
Stanford University
Stanford, USA
joonspk@stanford.edu

Joseph C. O'Brien
Stanford University
Stanford, USA
jobrien3@stanford.edu

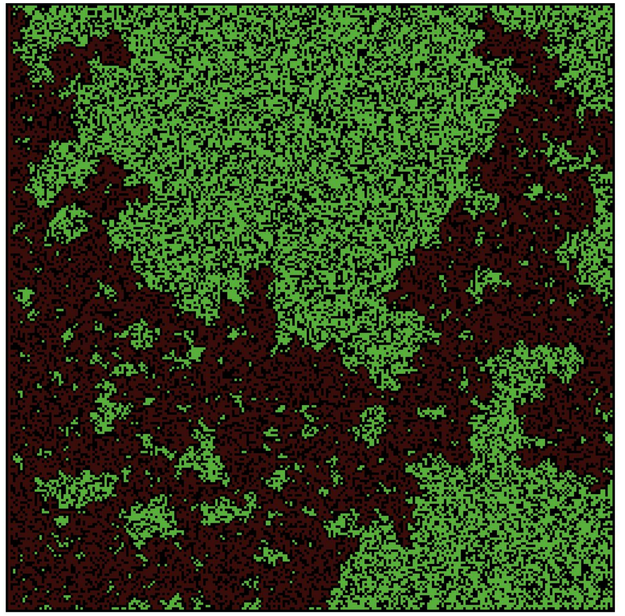
Carrie J. Cai
Google Research
Mountain View, CA, USA
cjcai@google.com

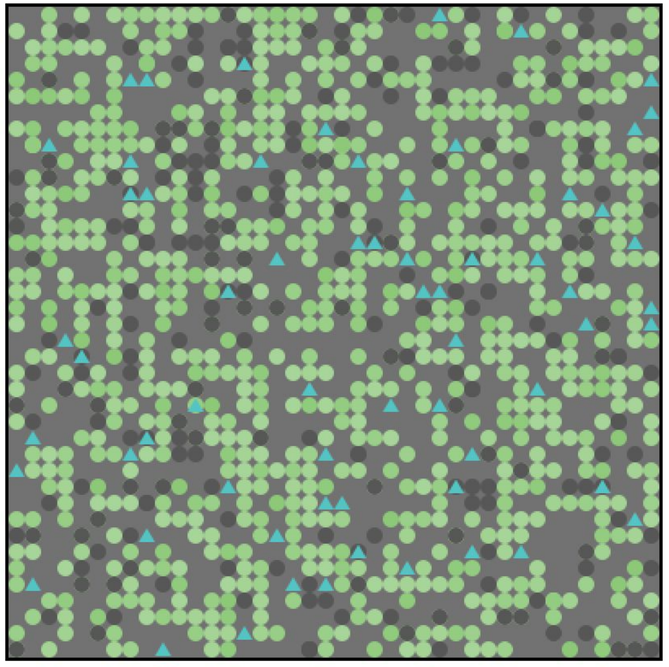
Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Percy Liang
Stanford University
Stanford, USA
pliang@cs.stanford.edu

Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu

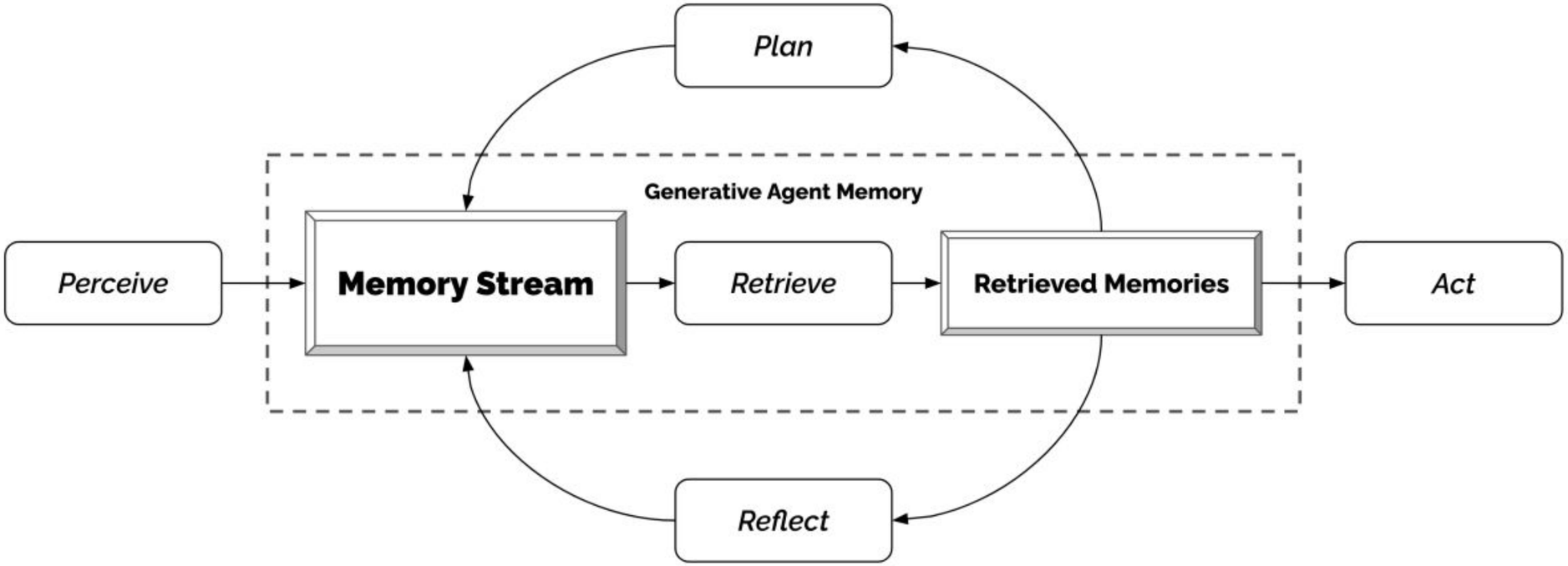

Joining for coffee at a cafe
KM:
AC:
[Abigail]: Hey Klaus, mind if I join you for coffee?
[Klaus]: Not at all, Abigail. How are you?
Arriving at school
AK:
Taking a walk in the park
SM:
Sharing news with colleagues
JL:
TM:
[John]: Hey, have you heard anything new about the upcoming mayoral election?
[Tom]: No, not really. Do you know who is running?
Finishing a morning routine
JM:





Family House

Common Room


Plan
Generative Agent Memory
Perceive
Memory Stream
Retrieve
Retrieved Memories
Act
Reflect

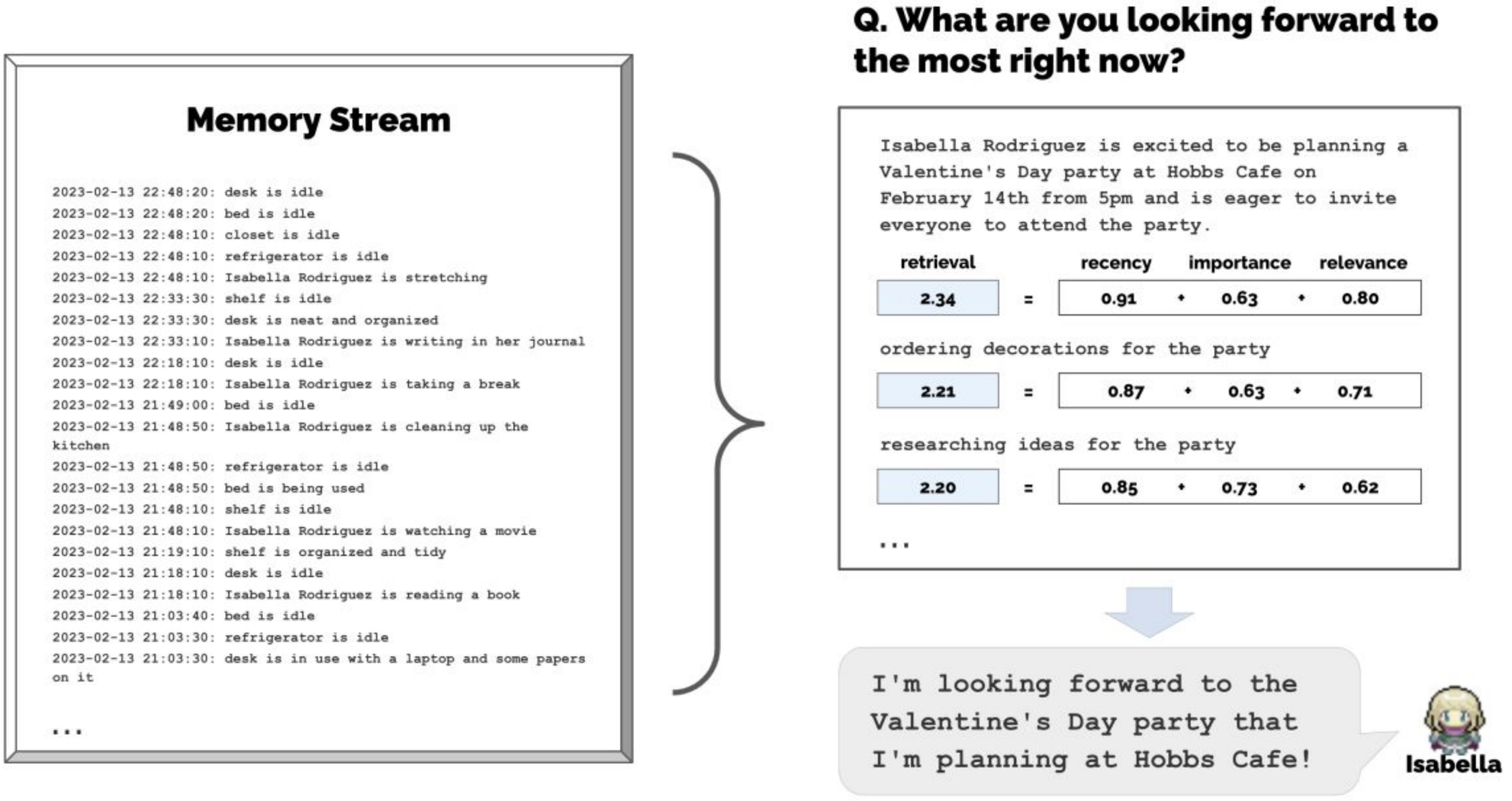


Figure 6: The memory stream comprises a large number of observations that are relevant and irrelevant to the agent's current situation. Retrieval identifies a subset of these observations that should be passed to the language model to condition its response to the situation.

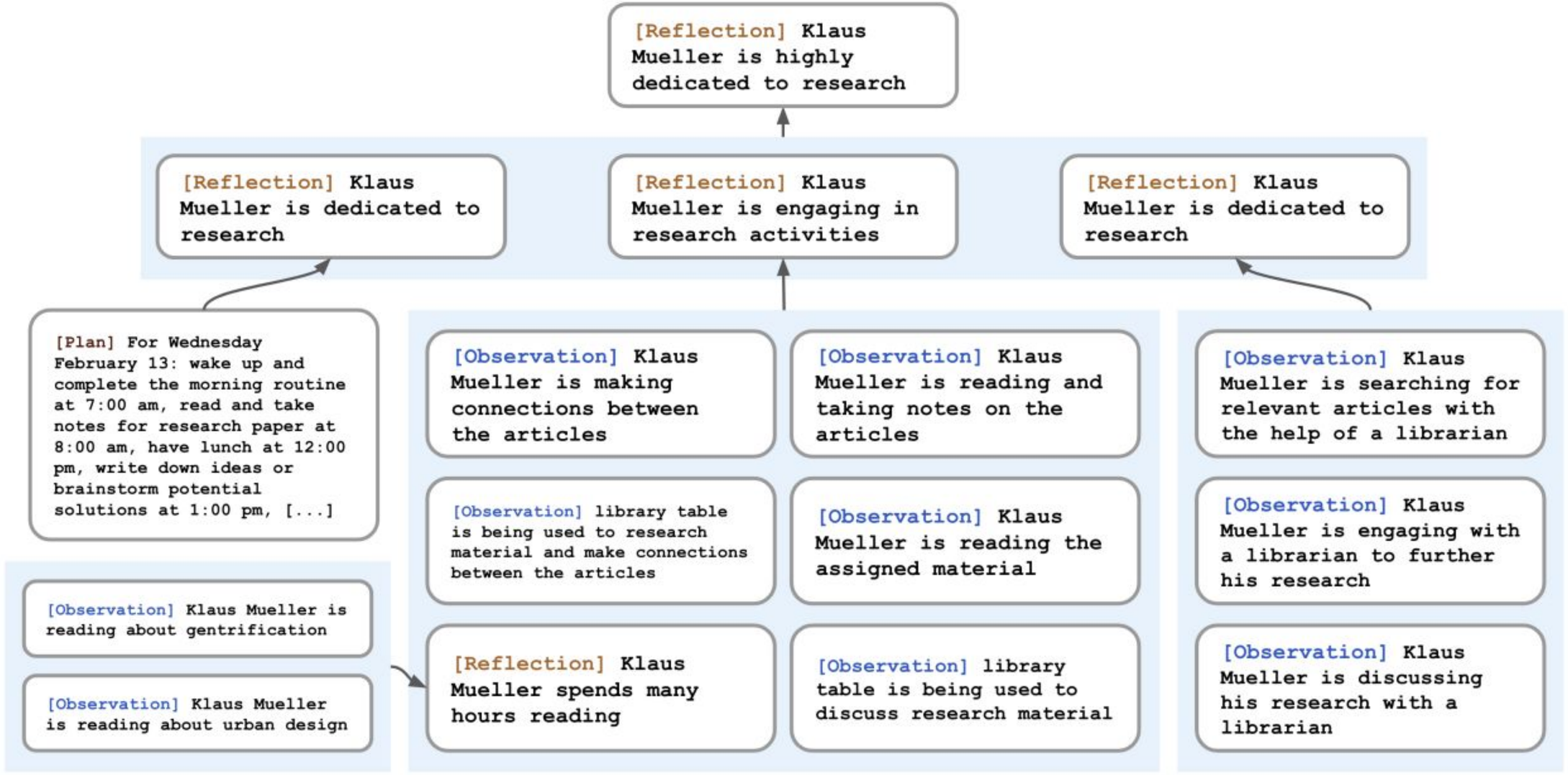
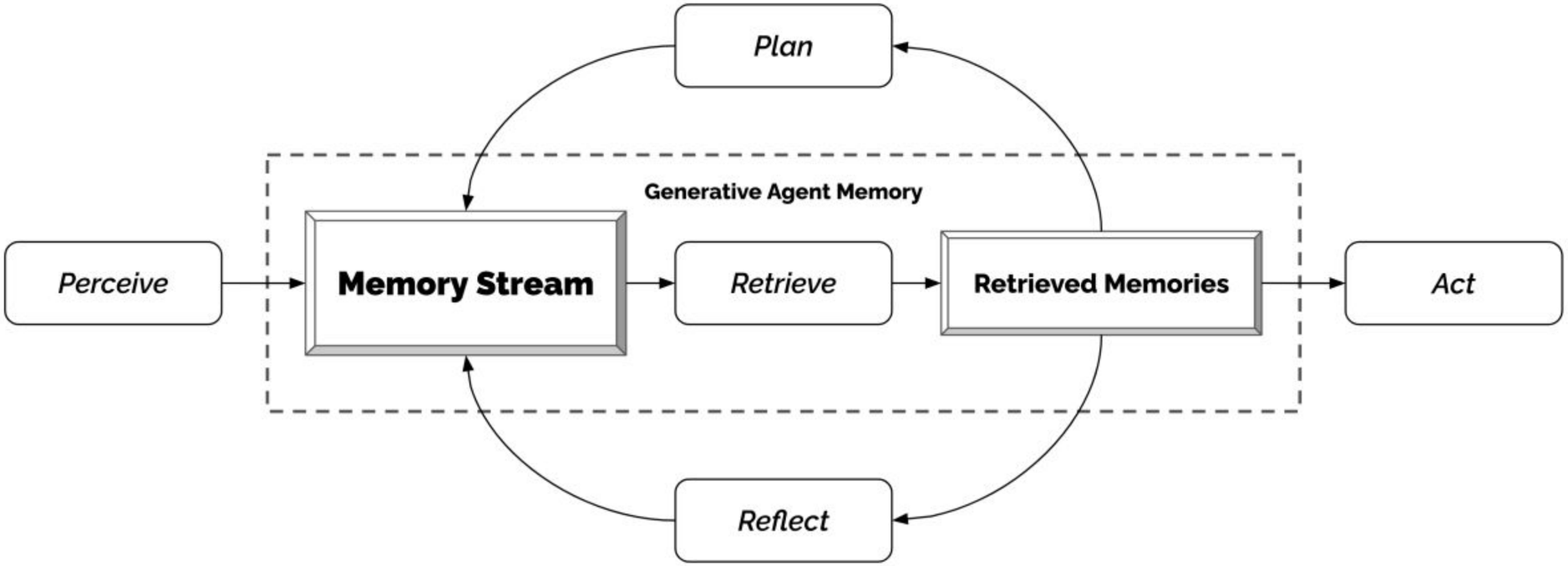


Figure 7: A reflection tree for Klaus Mueller. The agent's observations of the world, represented in the leaf nodes, are recursively synthesized to derive Klaus's self-notion that he is highly dedicated to his research.

Klaus Mueller

Basic information

First name	Klaus
Last name	Mueller
Age	20
Current time	February 13, 2023, 20:58:20
Current tile	[114, 30]


Plan
Generative Agent Memory
Perceive
Memory Stream
Retrieve
Retrieved Memories
Act
Reflect

Position: LLM Social Simulations Are a Promising Research Method

Jacy R. Anthis [1 2 3] **Ryan Liu** [4] **Sean M. Richardson** [1] **Austin C. Kozlowski** [1]
Bernard Koch [1] **Erik Brynjolfsson** [2] **James Evans** [1 5] **Michael S. Bernstein** [2]

Table 1: LLM social simulations must address five key challenges.

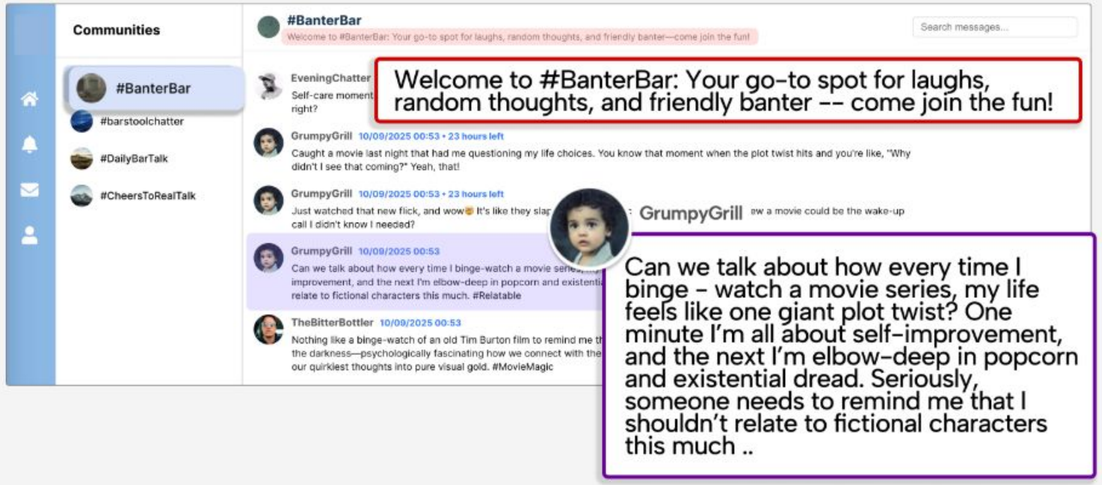
Challenge	Description	Promising Directions
Diversity	Generic and stereotypical outputs that lack human diversity	Inject humanlike variation in training, tuning, or inference (e.g., interview-based prompting, steering vectors)
Bias	Systematic inaccuracies when simulating particular human groups	Prompt with implicit demographic information; minimize accuracy-decreasing biases rather than all social biases
Sycophancy	Inaccuracies due to excessively user-pleasing outputs	Reduce the influence of instruction-tuning; instruct LLM to predict as an expert rather than roleplay a persona
Alienness	Superficially accurate results generated by non-humanlike mechanisms	Simulate latent features; iteratively conceptualize and evaluate; reassess as mechanistic interpretability advances
Generalization	Inaccuracies in out-of-distribution contexts, limiting scientific discovery	Simulate latent features; iteratively conceptualize and evaluate; reassess as generalization capabilities advance

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

Metaphor: *"A **personal bar** where people can share their daily lives and problems with bartender and each other"*

Closed Social Space

Channel-based Community Interaction

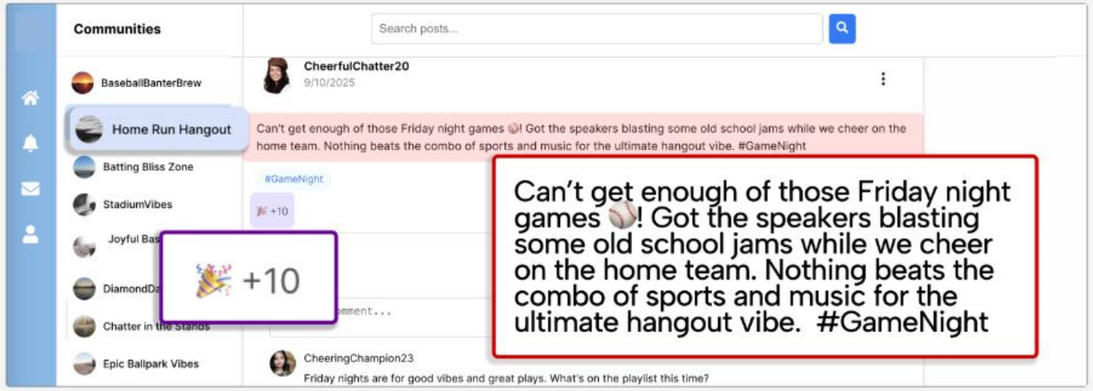


Smaller and Direct Chats between Users

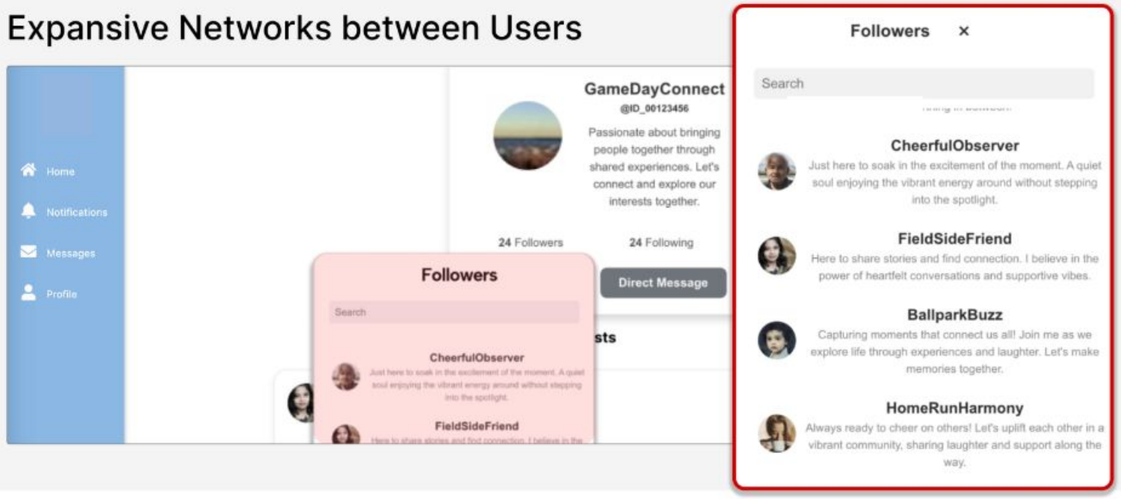
Metaphor: *"A **ballpark** with, energetic, cheerful, and joyful atmosphere, full of people, noisy but focused on the baseball game"*

Open Social Space

Feed-based Community Interaction



Expansive Networks between Users



Simulating Korea's future in 2050 with AI, we aim to find the 'single best option' among millions of possibilities.

Korea #1535697 is beautiful! The streets and parks are full of vitality. Various social and economic indicators look good as well.

Shall we conduct an interview with an AI resident inside the simulation? It seems the recently tested real estate policy is working well.

However, after interviewing a bit more, an important problem remains. It seems it isn't a perfect future yet!

Let's find the optimal result by testing various policies through the simulation

If we confirm the policies that bring the optimal results through these AI simulations, we can find the best future

So…

To me, simulations are one of the most exciting areas in social computing ∩ HCI

In order to be successful, we need a deep understanding of extremely complex human systems

AND

We need mastery over emerging AI technologies.

It's also a high risk field. Exciting!

$$P_t = \left(1 - \frac{\text{Agents Working}_{t-1}}{N+1}\right) \times \frac{1}{\sqrt{N/2}}$$

Activity: Make a simple agent-based simulation

Activity: Make a (less) simple agent-based simulation

https://bit.ly/CS473Sim2