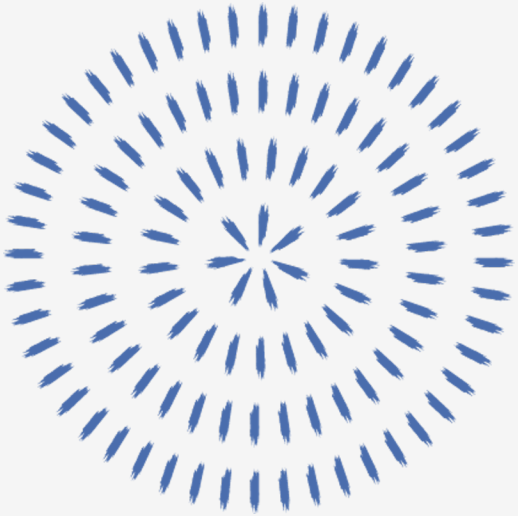
Paper Reading

Outlines

Data Augmentation for Simulating Interventions

Generative Agents



Data Augmentation

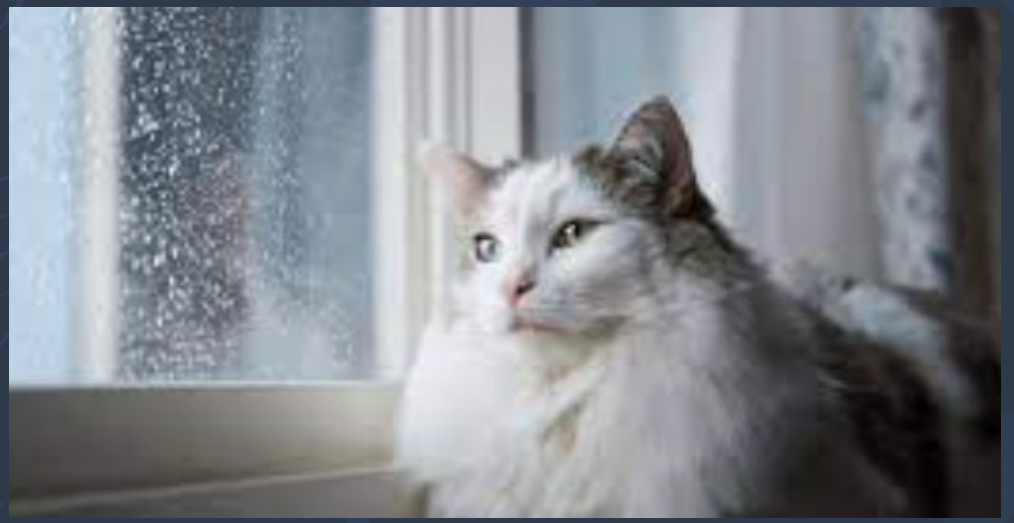
Abstract : observational data, generalization, spurious correlation, invariant features

Introduction

- Deep learning
- Domains
- Spurious correlation
  - cows & camels
- Methods -> lack of theoretical foundation(e.g)

Sample

- Domain adaption, enforce p(z|d=i) = p(z|d=j)
- E.g
  - Domain 1(晴天) vs Domain 2(雨天)
  - Taking photos of cats
- Problem : background caused the correlation.



Contribution Points

Intervention augmentation equivalence

algorithm -> select techniques

Conclusion

- SCM, address the core problem
  - domain generalization
- Impractical -> observational data
- Algorithm

Method

Goal : domain generalization, ML -> generate well to unseen domains

Domain

e,g rainy, sunny, cloudy.....

Scope of dataset

assume, domain has spurious correlation with the target(indeed, it did.)

training(N domains)

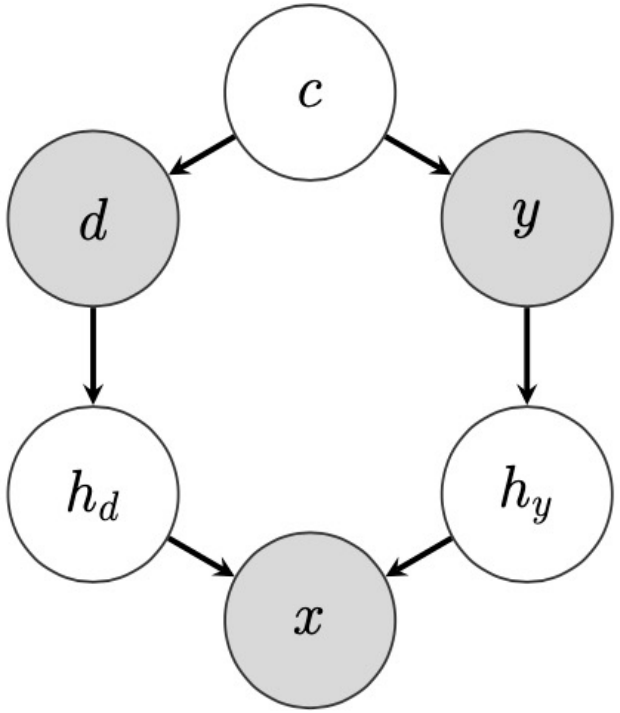
testing(include an extra unseen domain)

From causal perspective

- Samples
  - Country, hidden confounder
  - affect animal type and landscape(spurious correlation)

From causal perspective

- SCM(Structural Causal Models)
  - backdoor path(from hidden confounder)
  - d <- c -> y,
    - makes d is correlated to y, but it's non-causal.



$$d := f_D(c)$$
$$y := f_Y(c)$$
$$h_d := f_{H_d}(d)$$
$$h_y := f_{H_y}(y)$$
$$x := f_X(h_d, h_y), \quad (1)$$

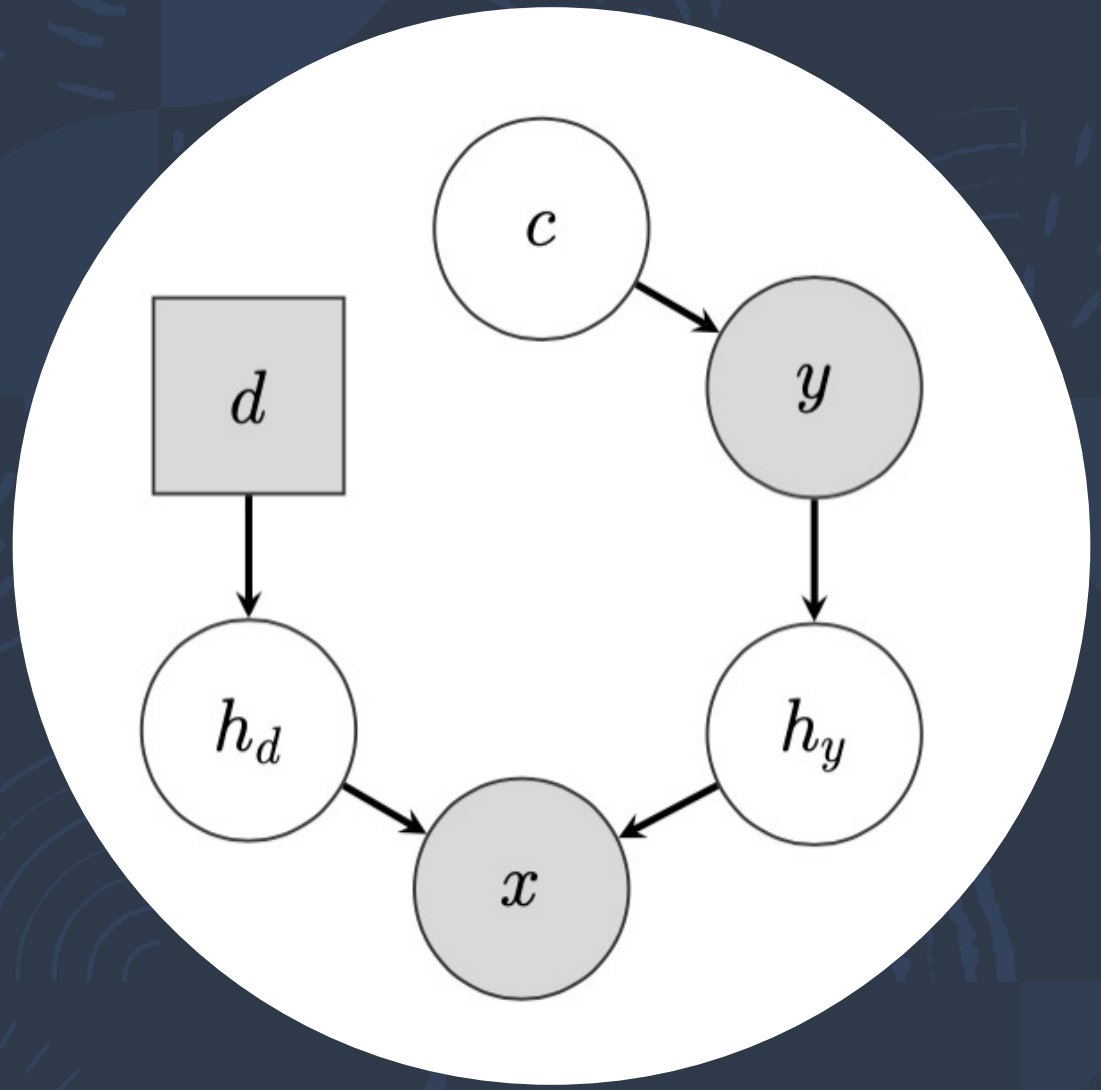
Figure 1. DAG and SCM with a hidden confounder.

Simulating Interventions

- Intervention on the domains
- Intervention on the domains' children
- Works for some cases
- Selecting algorithm

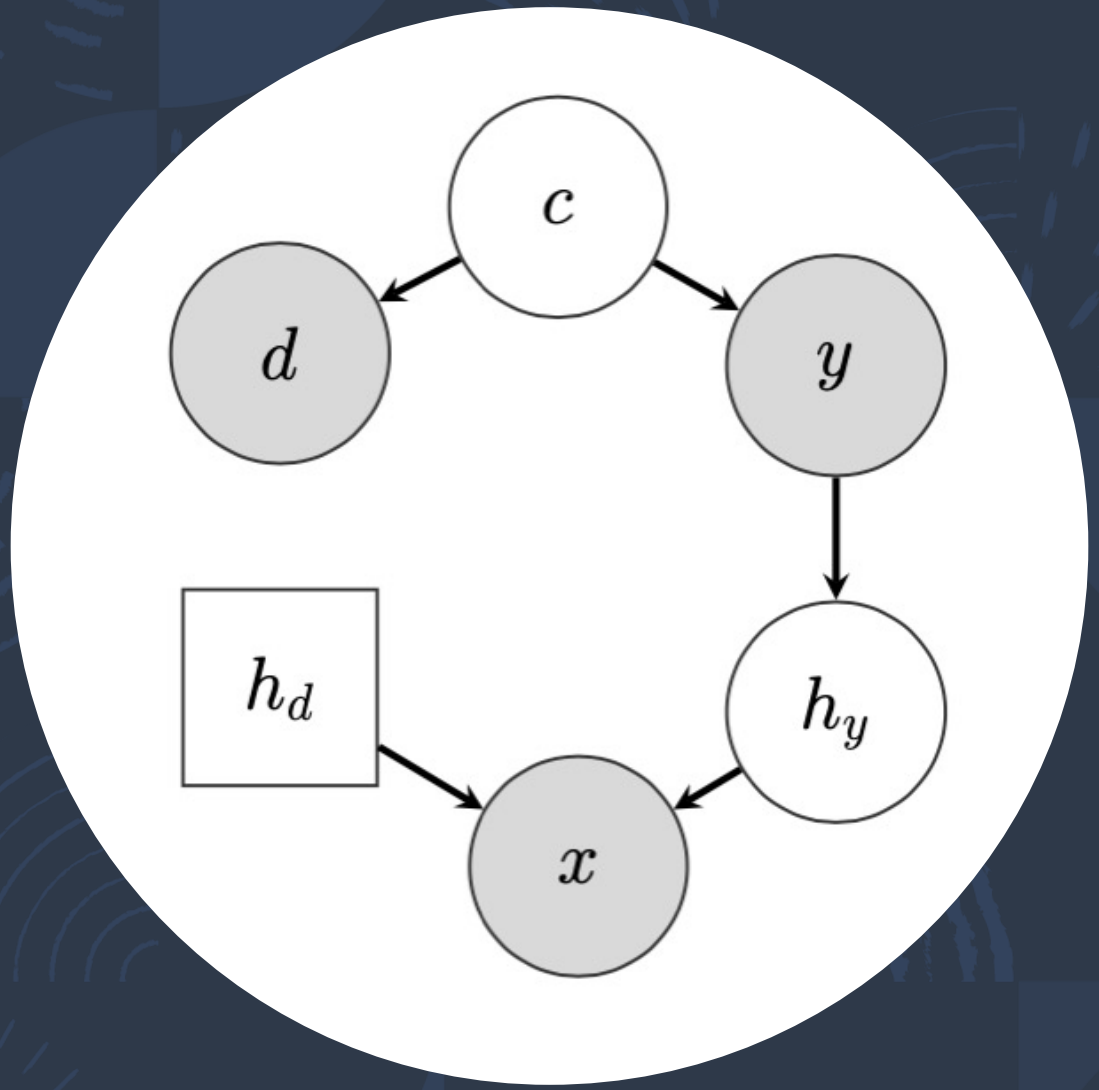
Simulating Interventions

- Intervention on the domains
  - background -> cow & camels
  - not feasible when collecting the data
    - observational data



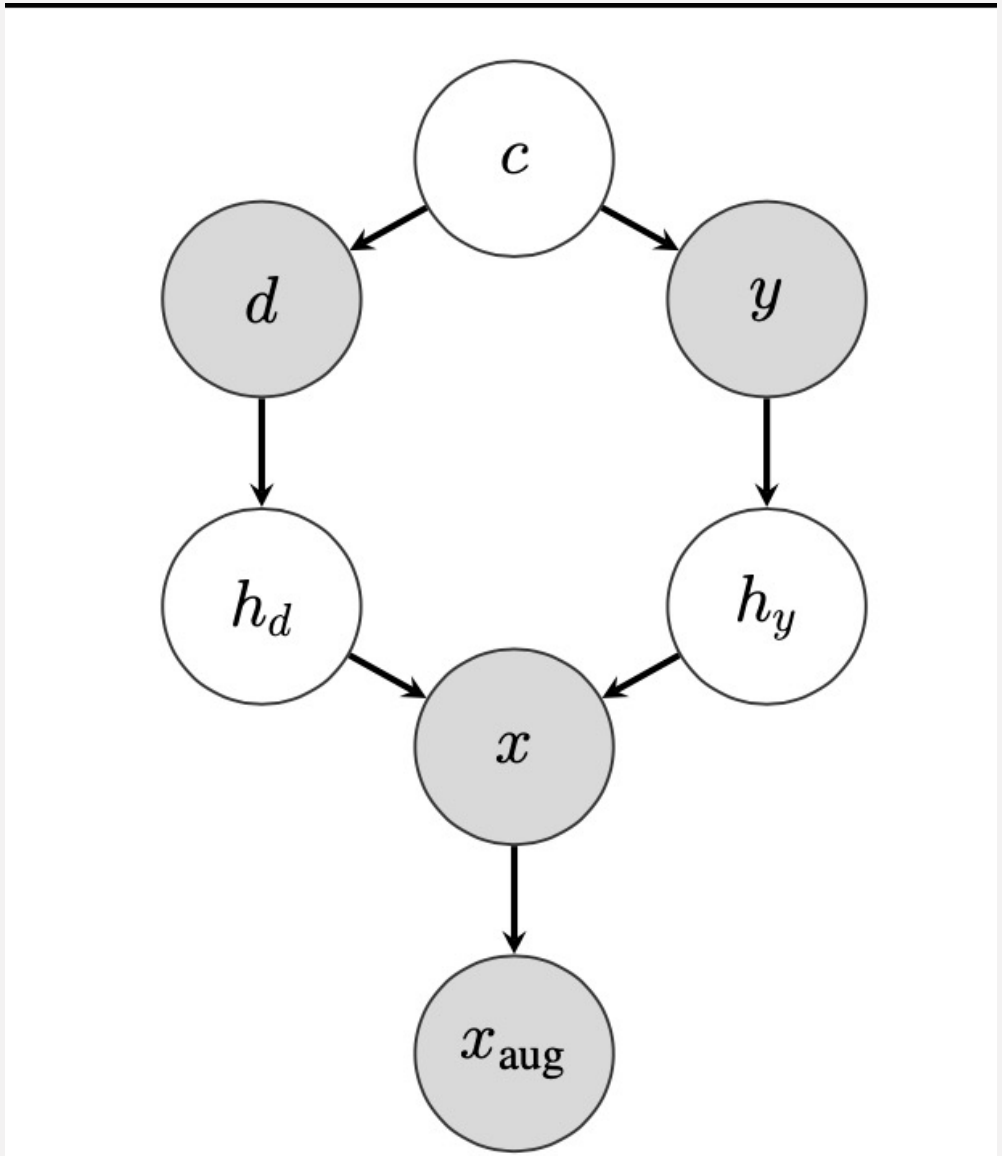
Simulating Interventions

- Intervention on the domains' children
  - Background includes : colors, intensity,....etc
  - Not feasible either for same reason



Simulating Interventions

- Some cases may work
  - data augmentation in combination with observational data
  - Assume target y & feature h_y, Unchanged
  - Fake Intervention



Simulating Interventions

- Samples
  - perturb the color inside the dataset
  - e.g, taking photo of one product, which is red. When you perturb the color of the image, it could represent as blue, yellow. Make sure, the product in real world align with the same model.

- One principle
  - Xaug = aug(x)
  - Intervention-augmentation equivalence
  - Limited cases.

Evaluation

- Theoretical -> Practical ?(Small datasets)
- Process of generalization
  - Ignore details
  - More focus on how to train a good model
  - As the regulation and environment better, we can make a good use of combination of both causal & techs.


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?
Taking a walk in the park
SM:
Arriving at school
AK:
Generative Agents
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:

The Sims vs Generative Agents

Objectives and Play Differences

- virtual experience
- generating new data, content, or decisions

Autonomy and creativity

- limited by the rules and design of the game
- Independent decision-making and creativity

Abstract

- Believerable Agents -> Human behaviour
- Large Language Model -> store and retrieve records
- Agents architecture
  - observation
  - planning
  - reflection

Introduction

- The sims
- Envision for over 4 decades.
- Problems
- workflow
- LLM & Human Behavior

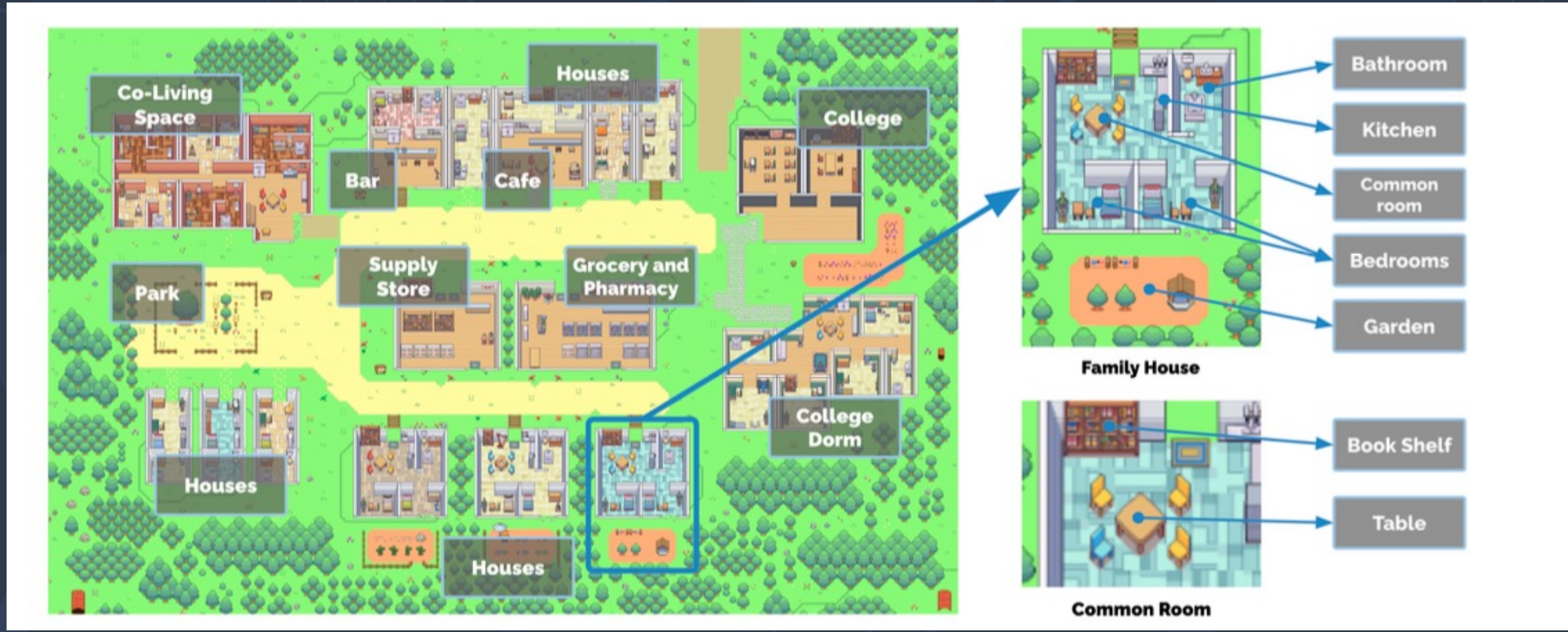
LLM & Human Behavior

- prompt chain
  - generate short natural language description of personas and their behavior
- LLM is key gradient
  - updated at each time step
  - reinforce
  - contradict

Conclusion

- Generative Agent
  - Deepening its understanding of itself
  - Non-player
  - immersive environments

Agent Behavior and Interaction

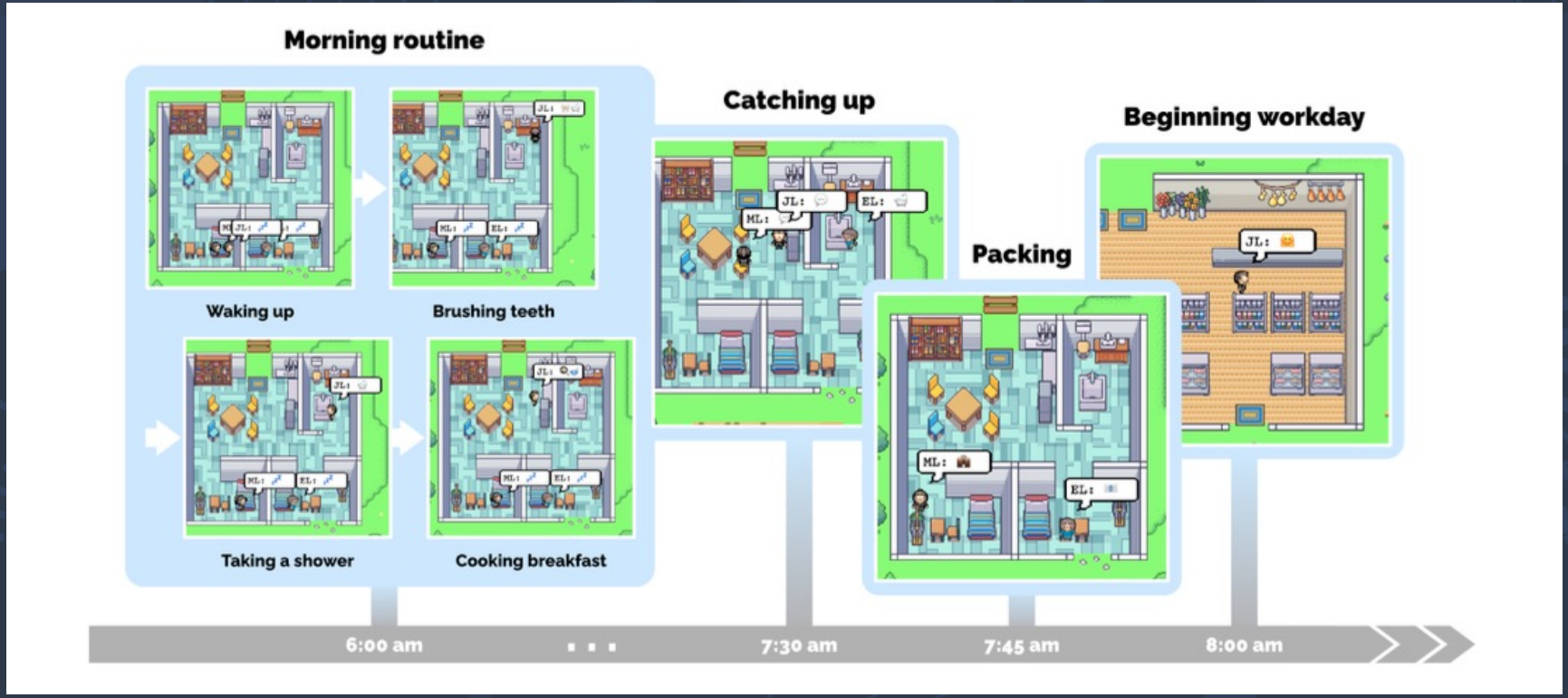


Communication

Sample

John Lin is a pharmacy shopkeeper at the Willow Market and Pharmacy who loves to help people. He is always looking for ways to make the process of getting medication easier for his customers; John Lin is living with his wife, Mei Lin, who is a college professor, and son, Eddy Lin, who is a student studying music theory; John Lin loves his family very much; John Lin has known the old couple next-door, Sam Moore and Jennifer Moore, for a few years; John Lin thinks Sam Moore is a kind and nice man; John Lin knows his neighbor, Yuriko Yamamoto, well; John Lin knows of his neighbors, Tamara Taylor and Carmen Ortiz, but has not met them before; John Lin and Tom Moreno are colleagues at The Willows Market and Pharmacy; John Lin and Tom Moreno are friends and like to discuss local politics together; John Lin knows the Moreno family somewhat well — the husband Tom Moreno and the wife Jane Moreno.

Sample


Morning routine
Waking up
Brushing teeth
Taking a shower
Cooking breakfast
Catching up
Packing
Beginning workday
6:00 am
7:30 am
7:45 am
8:00 am

- aware of other agents in local area
- communication mechanism & Learning

Sample – Inter-agent comm

John: My friends Yuriko, Tom and I have been talking about the upcoming election and discussing the candidate Sam Moore. We have all agreed to vote for him because we like his platform.

Sample – User Controls

Sample – Environment Interacton

- User Enter
- State of objects

Sample

- Emergent Social Behaviors
  - Information diffusion

Sam: Hey Tom, how's it going?
Tom: Good, thanks. What's up?
Sam: Well, I wanted to talk to you about something. I'm actually running for mayor in the upcoming local election.
Tom: Really? That's great news! Why are you running?
Sam: I've been involved in local politics for years now,

John: I heard that Sam Moore is running for mayor in the local election. Do you think he has a good chance of winning?
Tom: I do think he has a good chance. He's been working hard in the community and I think he will get a lot of

start. While taking a walk in Johnson Park, Sam runs into Latoya, and they introduce themselves and Latoya mentions that she is working on a photography project: "I'm here to take some photos for a project I'm working on." In a later interaction, Sam's interactions with Latoya indicate a memory of that interaction, as he asks "Hi, Latoya. How is your project going?" and she replies "Hi, Sam. It's going well!"

Sample - Relationship Memory

Form new relationship

Sample - Coordination

Generative Agent Architecture

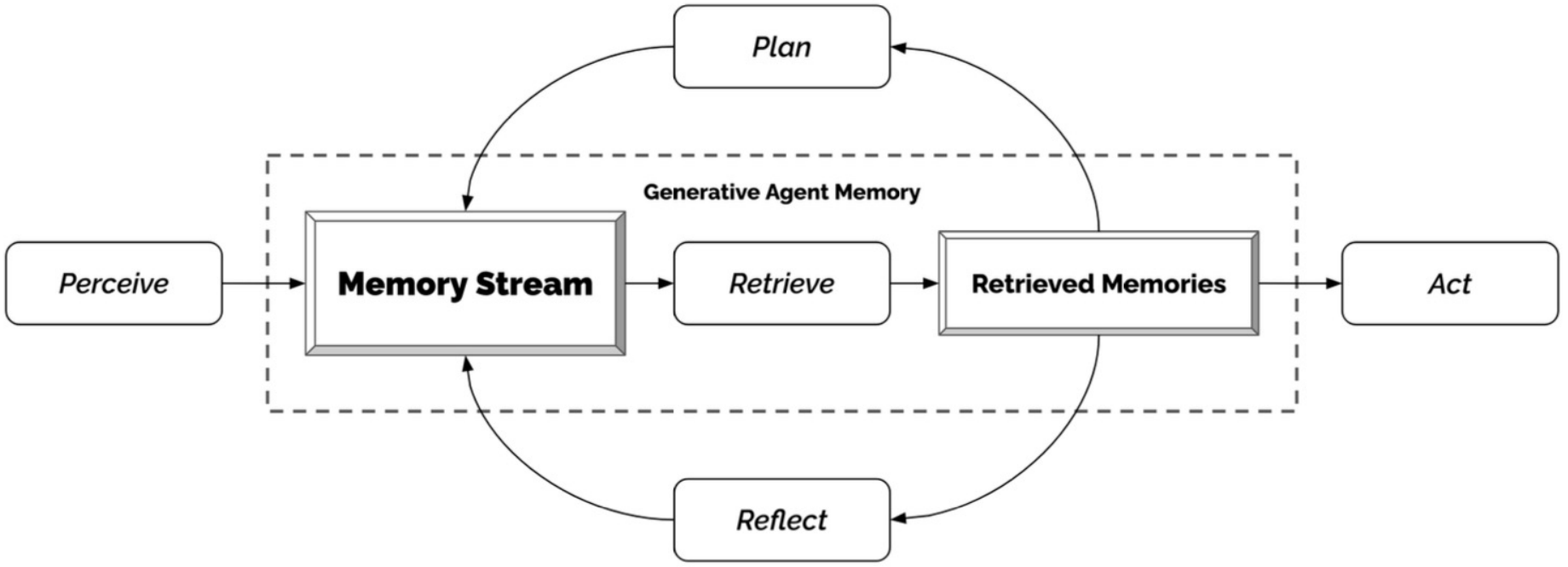
Framework

- Input :
  - current environment
  - past experiences
- Output :
  - behavior

Challenges :

- long-term planning
- coherence remain
- most relevant memory are retrieved and synthesized

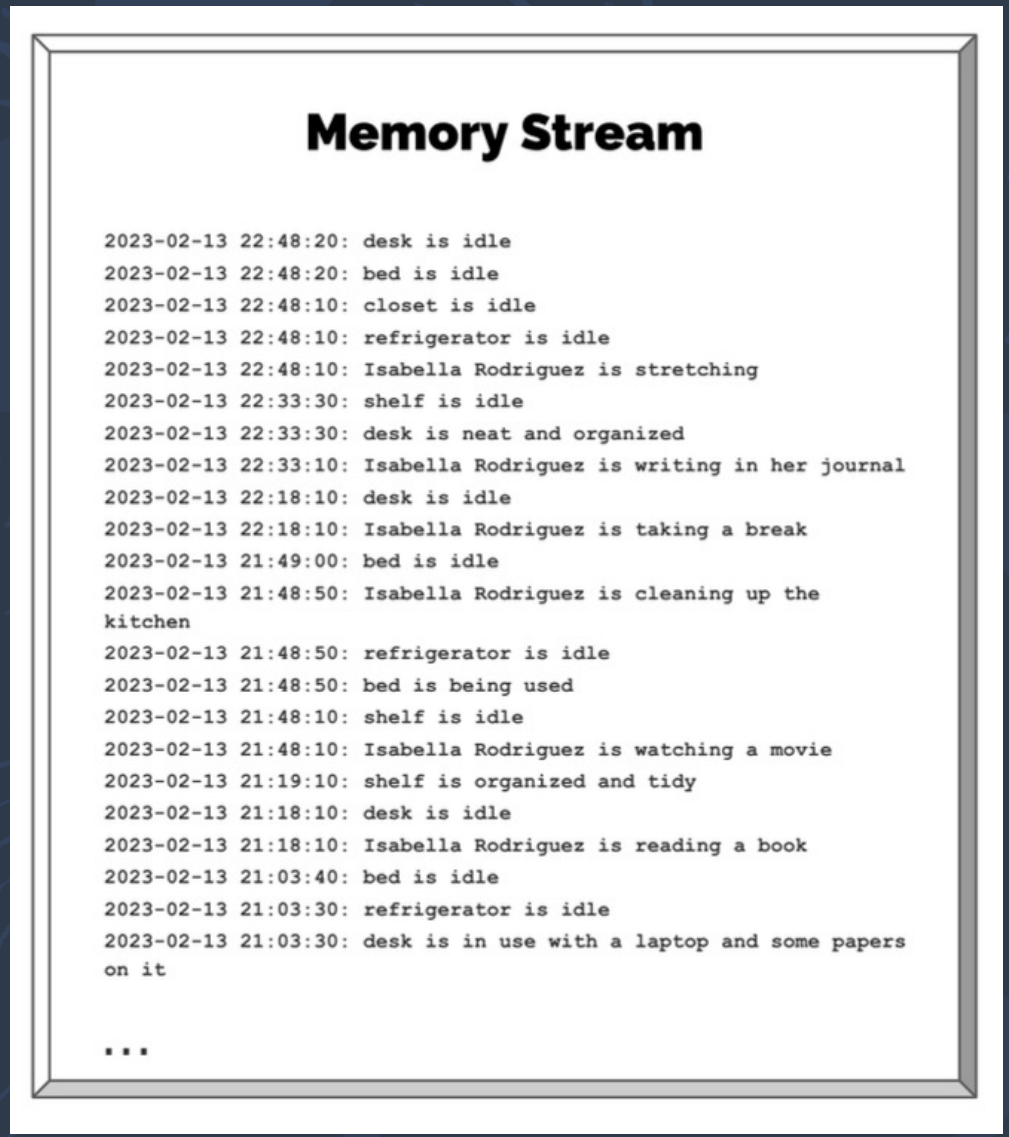
Architecture



Memory & Retrieval

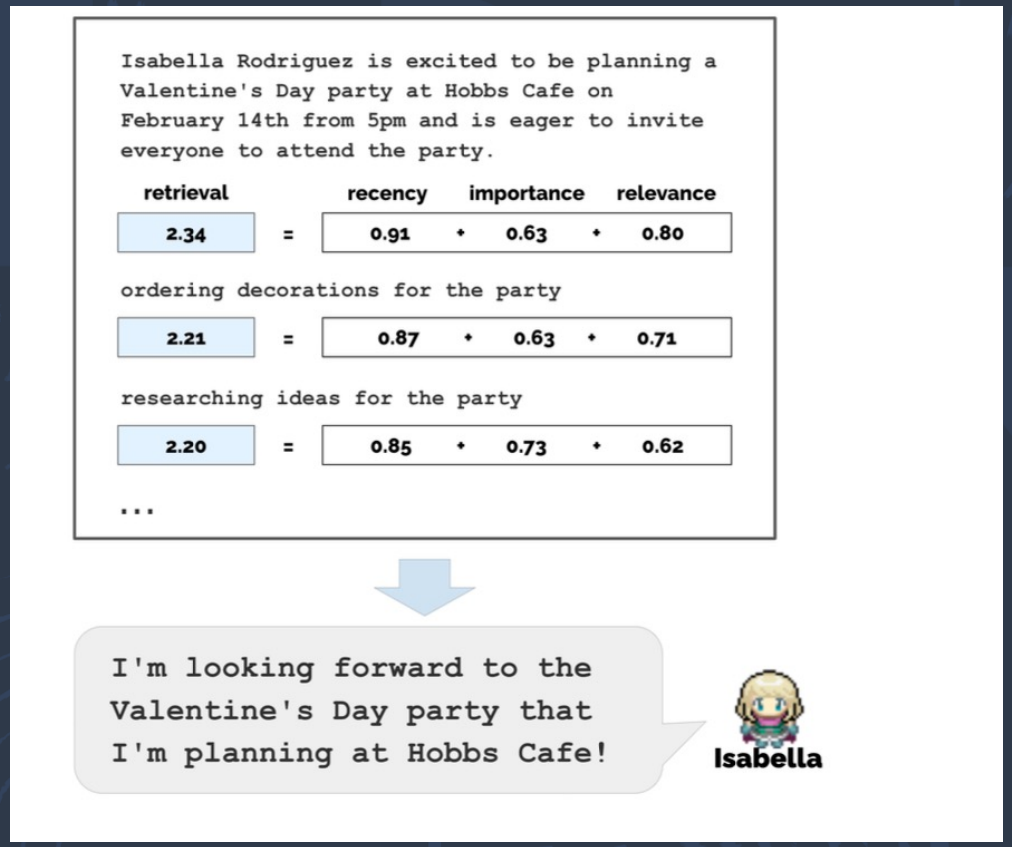
Challenges

- larger compared to the prompt
- Context window



Approach

- Observation
  - behaviors
- Retrieval function
  - input, current agent's situation
  - output, subsets of memory stream
  - pass the output to language model



Retrieval function

- Recency
- Importance
- Relevance

Recency

Recency assigns a higher score to memory objects that were recently accessed

decay factor is 0.99

Importance

signing a higher score to those memory objects that the agent believes to be important(using prompt model)

```
On the scale of 1 to 10, where 1 is purely mundane
(e.g., brushing teeth, making bed) and 10 is
extremely poignant (e.g., a break up, college
acceptance), rate the likely poignancy of the
following piece of memory.
Memory: buying groceries at The Willows Market
and Pharmacy
Rating: <fill in>
```

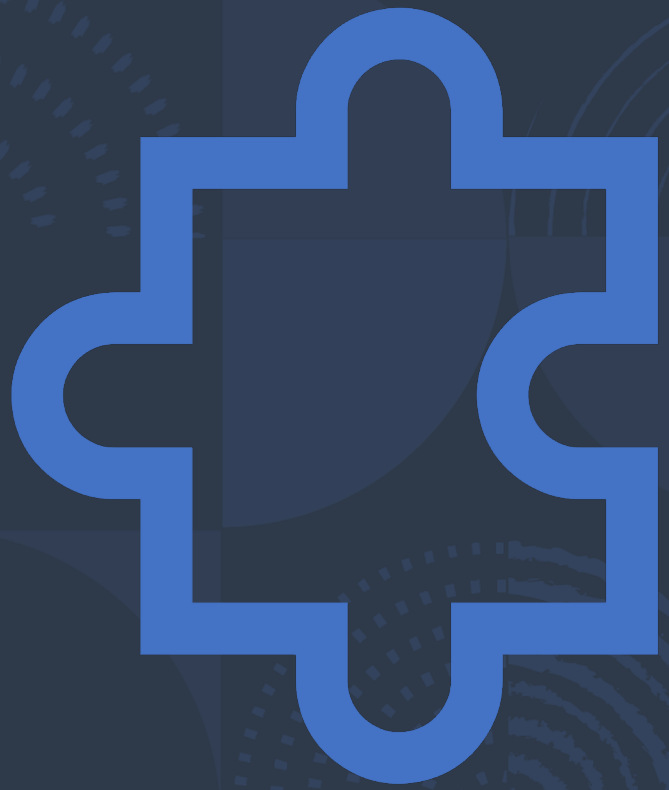
Relevance

- assigns a higher score to memory objects that are related to the current situation
- memory
  - generate description as embedding vector
  - retrieval memory as embedding vector
  - cosine similarity -> relavance

Reflection

Challenges

- struggle to generalize or make inferences
- "If you had to choose one person you know to spend an hour with, who would you choose?"
  - General Agent vs High-level Agent

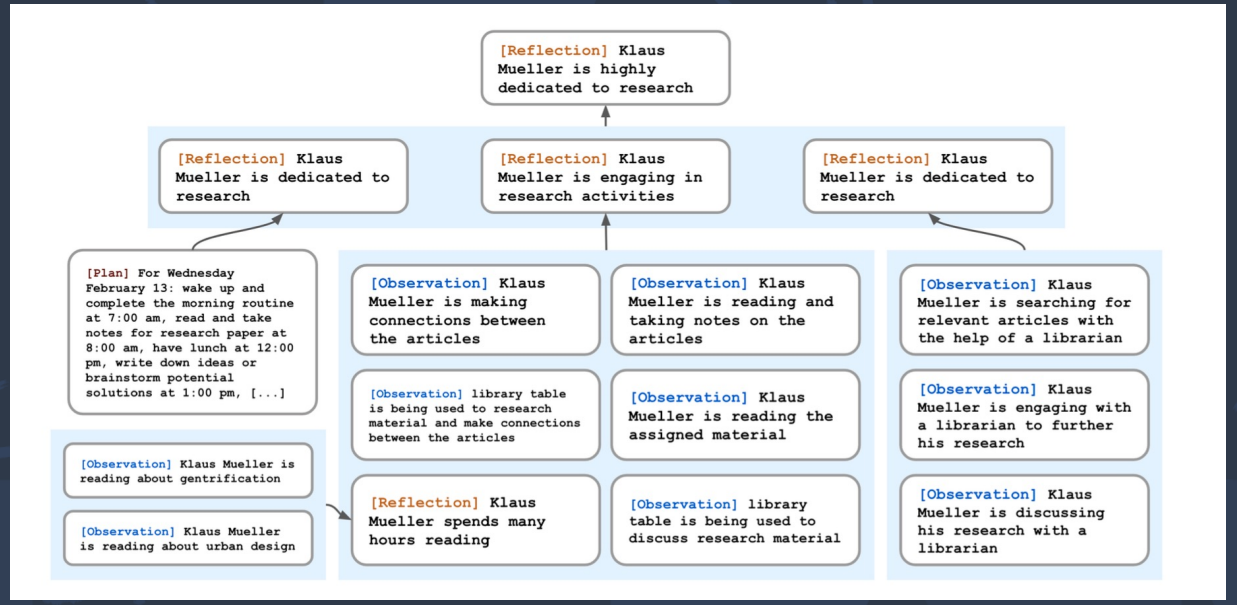


Approach

- a second type of memory
- generated periodically
  - importance scores exceeds a certain threshold.(150)

Sample

- Query LLM with 100 most recent records
- Prompt the language model
  - Get 3 most salient high-level questions
- Use the questions to retrieval from memory stream
- Prompt the language model
  - extract insights.



Statements about Klaus Mueller
1. Klaus Mueller is writing a research paper
2. Klaus Mueller enjoys reading a book on gentrification
3. Klaus Mueller is conversing with Ayesha Khan about exercising [...]
What 5 high-level insights can you infer from the above statements? (example format: insight (because of 1, 5, 3))

Planning & Reacting

- Challenge:
  - plan over a longer time horizon to ensure that their sequence of actions is coherent and believable
  - general prompt language model
  - high-level model output

Approach

- keep the agent's behavior consistent over time
  - a location, a starting time, and a duration
- plans are stored in the memory stream
  - Agents may change their plans midstream if needed

Sample

Name: Eddy Lin (age: 19)
Innate traits: friendly, outgoing, hospitable
Eddy Lin is a student at Oak Hill College studying music theory and composition. He loves to explore different musical styles and is always looking for ways to expand his knowledge. Eddy Lin is working on a composition project for his college class. He is taking classes to learn more about music theory. Eddy Lin is excited about the new composition he is working on but he wants to dedicate more hours in the day to work on it in the coming days
On Tuesday February 12, Eddy 1) woke up and completed the morning routine at 7:00 am, [. . .] 6) got ready to sleep around 10 pm.
Today is Wednesday February 13. Here is Eddy's plan today in broad strokes: 1)

[Agent's Summary Description]
It is February 13, 2023, 4:56 pm.
John Lin's status: John is back home early from work.
Observation: John saw Eddy taking a short walk around his workplace.
Summary of relevant context from John's memory: Eddy Lin is John's Lin's son. Eddy Lin has been working on a music composition for his class. Eddy Lin likes to walk around the garden when he is thinking about or listening to music.
Should John react to the observation, and if so, what would be an appropriate reaction?

Reacting and updating plans

- an action loop
- perceived observations are stored in their memory stream
- prompt and decide
  - continue or react to change

John Lin's status: John is back home early from work.
Observation: John saw Eddy taking a short walk around his workplace.
Summary of relevant context from John's memory: Eddy Lin is John's Lin's son. Eddy Lin has been working on a music composition for his class. Eddy Lin likes to walk around the garden when he is thinking about or listening to music.
John is asking Eddy about his music composition project. What would he say to Eddy?

Dialogue

- generate agents' dialogue by conditioning their utterances on their memories
- Result
  - "Hey Eddy, how's the music composition project for your class coming along?"

Discussion & Future work

- Act as proxy
- Improve retrieval function
  - more information generated
- Improve architecture’s performance
  - cost-effective(2 days simulation but cost thousands of dollars in token credits and taking multiple days to complete)
- Parallelizing agents
- Struggle to generate believable behavior for certain subpopulations
- Human emotions ; never be a substitute for real human input
- Error
  - wrong Inference to user’s goal
  - misinformation generation

Evaluation & Open Talk