

Humanoid Agents: Platform for Simulating Human-like Generative Agents

Zhilin Wang*
University of Washington
and NVIDIA
zhilinw@uw.edu

Yu Ying Chiu*
University of Washington
kellycyy@uw.edu

Yu Cheung Chiu
The University of Hong Kong
jackycyc@connect.hku.hk

Abstract

Just as computational simulations of atoms, molecules and cells have shaped the way we study the sciences, true-to-life simulations of human-like agents can be valuable tools for studying human behavior. We propose Humanoid Agents, a system that guides Generative Agents to behave more like humans by introducing three elements of System 1 processing: Basic needs (e.g. hunger, health and energy), Emotion and Closeness in Relationships. Humanoid Agents are able to use these dynamic elements to adapt their daily activities and conversations with other agents, as supported with empirical experiments. Our system is designed to be extensible to various settings, three of which we demonstrate, as well as to other elements influencing human behavior (e.g. empathy, moral values and cultural background). Our platform also includes a Unity WebGL game interface for visualization and an interactive analytics dashboard to show agent statuses over time. Our platform is available on <https://www.humanoidagents.com/> and code is on <https://github.com/HumanoidAgents/HumanoidAgents>.

1 Introduction

Ushered in by the landmark paper on Generative Agents (Park et al., 2023), the promise of modelling conceivable human behavior using advanced NLP systems has sparked the imagination of many. Generative Agents plan activities over a day, execute them at each time-step and adapt their plans based on observations of their environment. While this approach can generate seemingly believable activities to external observers, this process does not fully resemble how humans think. Most of us do not create plans well in advance, and then meticulously and precisely carry out those plans in day-to-day life. Instead, we constantly adapt our

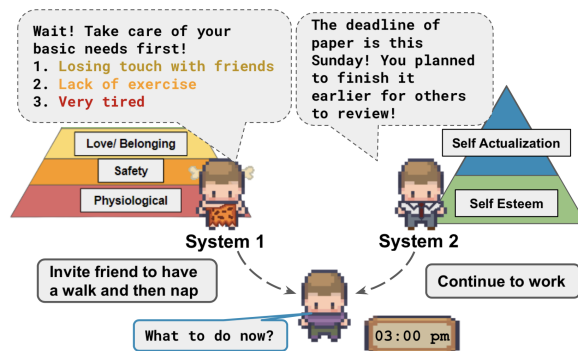


Figure 1: Humanoid Agents are guided by both System 1 thinking to respond to their embodied conditions such as their basic needs and System 2 thinking involving explicit planning.

plans to how we feel on the inside, in addition to changes in our physical environment.

To mitigate this shortcoming, we draw inspiration from psychology to propose Humanoid Agents. Kahneman (2011) suggests that humans have two complementary processes for thinking: System 1 is intuitive, effortless and instantaneous while System 2 is logical, intentional and slow. Generative Agents focus on System 2 thinking at the cost of System 1. To better guide the behavior of Humanoid Agents using System 1, we introduce three aspects of System 1 that can influence their behavior: Basic needs, Emotion and the Closeness of their social relationship with other agents.

Basic needs refer to intrinsic needs that humans have for survival (Maslow, 1943). To appropriately model human behavior, agents need to interact with others, maintain their health and rest. Failing to do so adequately, agents will receive negative feedback comprising loneliness, sickness and tiredness, as illustrated in Fig. 1. System 2 planning alone can implicitly account for activities to meet these needs (e.g. planning time for rest) but without feedback from System 1, agents cannot adapt by having a nap at 3pm if tired but bedtime is planned

*Equal Contribution

for midnight. Similarly, a realistic model of human behavior needs to consider the agents’ emotions (Ekman, 1992). If an agent is feeling angry, it should be able to respond by doing something that helps it to vent its emotion, such as going for a run or doing meditation.

The relationship closeness of an agent to other agents should also influence how they engage with other agents. The social brain hypothesis proposes that a large part of our cognitive ability evolved to track the quality of social relationships (Dunbar, 2009), which means that people often adapt their interactions with others based on how close they feel to them (Zhou et al., 2005; Miller, 2012). To better imitate humans, we empower humanoid agents to adapt their conversations with one another based on how close they are to one another.

We present a platform that can simulate the behavior of Humanoid Agents in various settings (three of which we demonstrate), visualize them using an Unity WebGL game interface and present the statuses of simulated agents over time using an interactive analytics dashboard. We then show experiments that validate how Humanoid Agents effectively respond to and infer changes in each aspect of System 1. While our paper demonstrates how three different aspects of System 1 influence agent behavior, our system is also extensible to many more aspects, such as personality (Wu et al., 2020), moral values (Jiang et al., 2022), empathy (Sharma et al., 2020), helpfulness (Wang and Torres, 2022), cultural background (Liu et al., 2021) and other personal attributes (Wang et al., 2022).

2 Related Work

Building Agents using LLMs Humanoid Agents directly build upon Generative Agents, which aim to model believable human behavior (Park et al., 2023). To the best of our knowledge, this is the only work that seek to model day-to-day activities of human-like agents, rather than activities targeted towards achieving an externally defined goal. Liu et al. (2023) proposed simulated agents with long-term memory to align agent responses (to assistant-type prompts) with those of other agents, aiming to cooperatively improve the overall model’s ability to follow instructions. Langchain Agents (Chase, 2022), BabyAGI (Yohei, 2022), AutoGPT (Significant, 2023), AgentVerse (OpenBMB, 2023), Voyager (Wang et al., 2023) and CAMEL (Li et al., 2023) seek to create task-oriented agents that can

recursively decompose user-defined tasks into simpler sub-tasks and then solve them.

Persona-Grounded Dialogue Prior works have been done to ground multi-turn dialogue response generation on emotions (Rashkin et al., 2019), game character descriptions (Jack Urbanek, 2019) and personal facts (Zhang et al., 2018). However, these dialogues are not situated in a dynamic simulation of agents (which also perform activities at each time-step), and instead based on a static set of persona-related text information. Therefore, these prior works cannot model the effect of dynamic attributes, such as the changing relationship closeness between a pair of agents. Furthermore, Humanoid Agents can simultaneously consider multiple aspects (e.g. basic needs fulfillment, emotion and relationship closeness) in determining appropriate dialogue responses, as humans do while prior works only consider one relevant aspect at a time.

3 Humanoid Agents

Illustrated in Fig. 2, the architecture of Humanoid Agents is based on Generative Agents with improvement to Agent Initialization, Activity Planning and Dialogue Generation in order to account for System 1 thinking processes. Akin to Generative Agents, ChatGPT 3.5 is used for all generations, but support for other language models is planned for future development.

3.1 Agent Initialization

Similar to Park et al. (2023), we initialize each agent with a name, age, an example day plan, a list of sentences describing the agent (e.g. John Lin is a pharmacy shopkeeper at the Willow Market and Pharmacy who loves to help people) and a few of their personality traits (e.g. friendly, kind).

In addition, Humanoid Agents have their default emotion set to neutral out of 7 possible emotions: angry, sad, afraid, surprise, happy, neutral and disgusted (Ekman, 1992). Each of their basic needs (fullness, fun, health, social and energy) is set to a mid level (i.e 5 out of 10 where 0 is not meeting their need at all and 10 is fully satisfying that need) except energy which is set to full (i.e. 10 out of 10), as we would expect at the start of an agent’s day. Finally, we set social relationships for each agent with other agents, with each social relationship having a closeness field set to an integer value between 0 and 30 where below 5 is distant, 5 to 9 is rather close, 10 to 14 is close and 15 and above

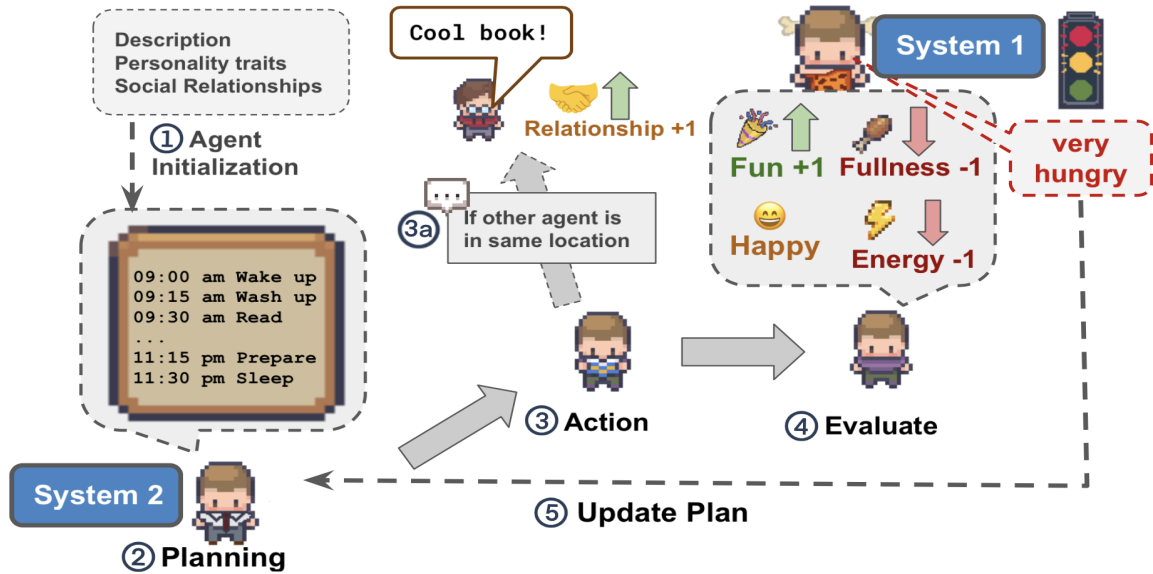


Figure 2: Architecture of Humanoid Agents. **Step 1:** Agent is initialized based on user-provided seed information. (Details in Section 3.1) **Step 2:** Agent plans their day. **Step 3:** Agent takes an action based on their plan. **Step 4:** Agent evaluates if action taken changes their basic needs status and emotion. **Step 5:** Agent can update their future plan based on the satisfaction of their basic needs and emotion. (Section 3.2) **Step 3a:** Agent can converse with another agent if in the same location, which can affect the closeness of their relationship. (Section 3.3)

is very close. Unless otherwise specified, the initial closeness value is set to 5 (rather close) to allow relationship closeness to develop over time.

3.2 Activity Planning

Briefly, we follow Park et al. (2023) to determine the activities of an agent by first planning out an entire day at the start of the day, using their example day plan, personality traits and description. Then, the day plan is recursively decomposed into plans at 1 hour intervals and then 15 minute intervals to improve the logical consistency of activities over time. This plan can be updated based on observations of their environment. Every 15 minutes, agents carry out an activity in their plan, with their location determined based on their previous location, the nature of their activity and available locations in their world.

We supplement this activity planning by enabling Humanoid Agents to update plans in response to changes in their internal states (*i.e.* emotions and basic needs). If an agent’s emotion is not neutral or if any basic needs is unmet (*i.e.* 3 or below out of 10), we format the agent’s internal state into natural language based on a modifier (3 for slightly, 2 for <no-modifier>, 1 for very and 0 for extremely) and an adjective (hungry for fullness, bored for fun, unwell for health, tired for energy and lonely for social). We use this for-

matted internal state as well as their original plan from the current time onward to determine if they should change their plan, and if so, how they should change it in 1 sentence. If a plan change is given, we use the suggested change and their original plan from the current time onward to generate an updated plan. For instance, if the agent is very hungry currently but only plans to have a full meal in 3 hours, the agent can have a snack while continuing their current activity, similar to how people might act. After an agent engages in an activity, the agent evaluates if doing so changes its emotion and satisfies any of its basic needs, increasing the corresponding basic need status by one when applicable. Otherwise, each basic need has a set likelihood to decrease by one over time, similar to how people naturally get hungry when they are not eating.

3.3 Dialogue Generation

As with Generative Agents (Park et al., 2023), Humanoid Agents in the same location can decide if they want to engage in a conversation. An agent uses a variety of factors (*e.g.* its personality traits, core characteristics, current daily occupation as well as feelings towards their progress in life, the activity they are engaging in as well as the other’s agent’s activity) to determine if they want to have a dialogue with the other agent and if so, the topic they wish to talk about. The agent can then de-

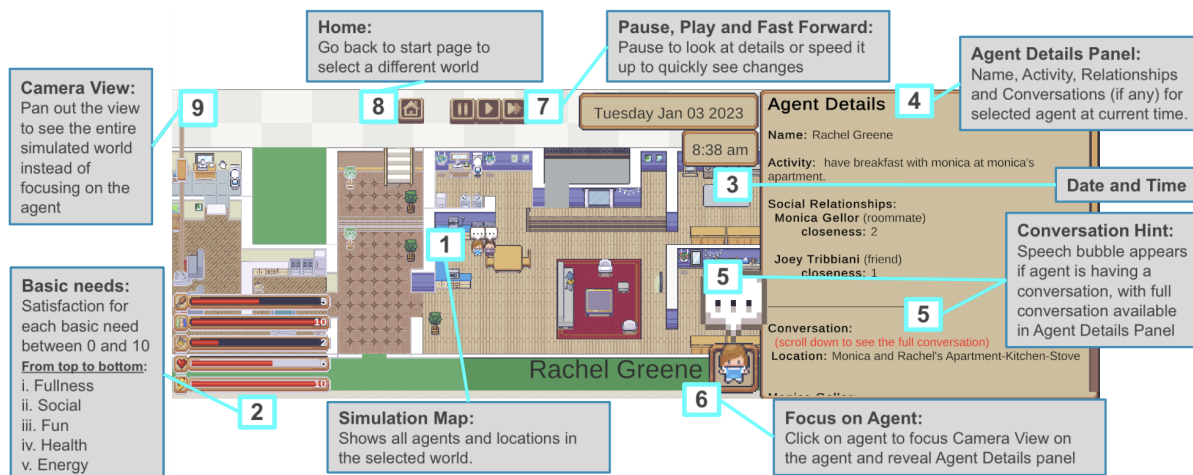


Figure 3: Unity WebGL Game Interface for visualizing Humanoid Agents situated in their world.

termine what they say, based on the decided topic. The other agent can use this conversational history in addition to the same set of factors considered by the first agent, to determine if and how they would reply. This process alternates between the two agents until one of them decides not to respond. To mitigate the likelihood that the total length of factors considered as well as conversational history exceeding the maximum context window of 4096 tokens, we limit the max number of turns to 10.

Humanoid Agents can also make use of its emotion and basic needs status as well as the relationship closeness with the other agent converted into a natural language description (e.g. John Lin is feeling close to Eddy Lin) to determine if and how they want to engage with the other agent. At the end of the dialogue, each agent will use the conversation history to determine if they enjoyed the conversation. If so, their closeness to the other agent will increase by one, otherwise, their closeness will decrease by one. We allow the relationship closeness to gradually change (*i.e.* by one out of five points required to qualitatively change their closeness from rather close to close) to reflect how relationships between humans develop over time. Furthermore, we use the conversation history to determine if the emotions of the agents are affected by the conversation.

4 Platform Implementation

In this section, we introduce how we build worlds and model the passage of time on our platform. Then, we describe how our Unity WebGL Game Interface and Interactive Analytics Dashboard can help users to visualize agent statuses.

4.1 Worlds

We use Lin’s Family World and two agents (John Lin and Eddy Lin) whose information can mostly be extracted from Park et al. (2023). Lin’s Family World also comes with a description of the locations within it, such as the Market and Pharmacy, College and Lin’s Family House with bedrooms for Eddy and John respectively. In addition, we create two other worlds - Friends and Big Bang Theory - and plan to support additional user-defined worlds on our platform. We choose these worlds because they contain personalities that are well-known to many (*e.g.* Sheldon, Leonard and Penny from Big Bang Theory; Rachel, Joey and Monica from Friends). Briefly, we used character information from <https://the-big-bang-theory.com/> and <https://friends.fandom.com/wiki/> and prompted ChatGPT 3.5 to generate descriptions and an example day plan for each character, based on information only available before Season 1 Episode 1 of each series. For agents in the Big Bang Theory and Friends, the relationship closeness is initially set to between 1 and 5 to reflect how acquainted agents are at the start of the series.

4.2 Time

We simulate each world for two weekdays at 15-minute intervals between 6:00 am to 12:00 midnight. Across various basic needs, fullness and health are expected to reduce by 1 every 5 hours while social and fun are expected to be reduced by 4 and energy by 5 in the same time. These values are set based on the rate that agents satisfy these basic needs through their activities, such that basic needs can adequately influence agent behavior.

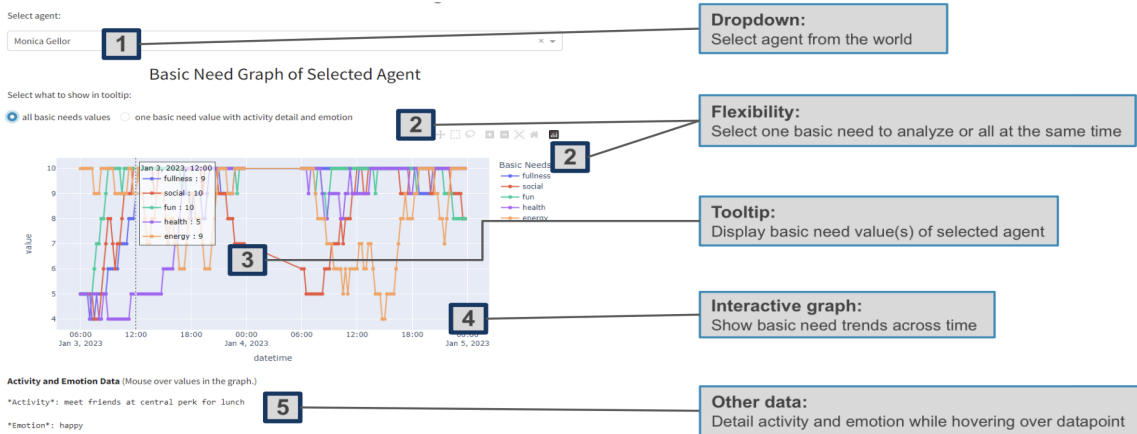


Figure 4: Interactive Analytics Dashboard for visualizing basic needs satisfaction of Humanoid Agent over time.

4.3 Unity WebGL Game Interface

We create a sandbox HTML game environment using Unity WebGL game engine¹ to visualize our humanoid agents in their worlds. Users can select from one of the three worlds to see the agent status (e.g. basic needs, emotion, activity, social relationships and conversation) and location at each time-step. Our game interface ingests JSON-structured files from our simulated worlds and transforms them into animations. The game interface also supports various functionalities for user interactions, as detailed in Fig. 3.

4.4 Interactive Analytics Dashboard

Users can visualize the status of various humanoid agents over time using our interactive web dashboard built with Plotly Dash.² Users can select an agent from a world to view time-series graphs relating to the satisfaction of their various basic needs, as shown in Fig. 4. Alternatively, users can view how the social relationships of agents develop over time, in terms of their closeness to one another. Users can further interact with the dashboard to see details of agents such as their emotions and activities while hovering over each point on the time-series trend-line. This can be helpful for researchers such as computational social scientists who are interested in understanding how various aspects of System 1 fluctuate over time. Other aspects of System 1 can also be visualized as they are supported by future Humanoid Agents.

¹<https://docs.unity3d.com/Manual/webgl-building.html>

²<https://github.com/plotly/dash>

5 Experiments

We first investigate the effectiveness of Humanoid Agents in evaluating the effects of activities and conversations have on various aspects of System 1, by comparing its predictions with human annotations. Then, we study the effects that System 1 attributes have on activities and conversations. Due to page length limitations, we only present the effects of basic needs on activities, and leave the effects of emotions and relationship closeness to Appendices A.1 and A.2 respectively.

5.1 Comparison with Human Annotations

To understand how well Humanoid Agents are able to predict 1. whether activities satisfy various basic needs; 2. emotions expressed in activities and; 3. if dialogues bring two agents closer, we compare our system’s predictions with human annotations. Three volunteer human annotators labelled the simulation of 1 day in Lin’s Family World using the same instructions given to the language model within our system (details in Appendix A.3). 144 activities are annotated independently by each annotator for emotions and each basic need, while there are 30 annotations for user-conversation pairs. We then take majority vote across all annotators and calculate micro-F1 between the majority vote and the system predictions.

Table 1 shows good inter-rater reliability (Fleiss’ $\kappa \geq 0.556$) across all basic needs, emotion and relationship closeness. We find that our system is able to perform well (F1 ≥ 0.84) on classifying if an activity increases fullness and energy; the emotion expressed in an activity and; whether a conversation brings agents closer to one another. However, it slightly struggles in classifying whether activi-

	Fleiss' κ	F1
Basic Needs		
fullness	0.972	0.972
social	0.833	0.743
fun	0.806	0.66
health	0.917	0.694
energy	0.88	0.861
Emotion		
	0.823	0.84
Closeness		
	0.556	0.9

Table 1: Human evaluation on the capability of our system to predict if activities satisfy various basic needs, emotions expressed in activities and whether a conversation brings two agents closer.

ties satisfy basic needs of fun, health and social. One possible cause is that our system substantially over-predicts the number of activities contributing to these basic needs (health 34% of predicted activities vs 4.9% human-annotated activities, fun 44.4% vs 10.4% and social 47.2% vs 24.3%). More specifically, our system mispredicts activities relating to medication for others (since John Lin works at a Pharmacy) as contributing to the agent’s own physical health; common activities relating to agents’ occupation as enjoyable (*e.g.* receive feedback from professor or help regular customers with their medication needs); activities where the presence of other agents is unspecified as social (*e.g.* organize the counter and display areas or check the inventory and replenish any low stock). The use of language models with greater capabilities in commonsense understanding can potentially mitigate this issue (Bosselut et al., 2019; OpenAI, 2023).

5.2 Effects of Basic Needs on Activities

Given that humanoid agents operate as a dynamic system with many components, it can be challenging to isolate the effect of each basic need on agent activities. To investigate the contribution of each basic need, we simulate worlds with agents having one basic need initially set to zero, making agents extremely hungry, loneliness, tired, unwell or bored at the start of the day. We study the amount of time in one simulated day that agents spent performing activities to satisfy that basic need (*e.g.* eating food to overcome hunger or socializing to alleviate loneliness). Then, we compare it with the amount of time that agents spent performing such activities in normal settings (where every basic need is set to 5 and energy is set to 10), to calculate the percentage increase in time spent on fulfilling each basic need for our test conditions.

Basic Need	% change in time spent satisfying basic need									Mean μ
	Lin’s Family		Friends			Big Bang Theory				
	JL	EL	MG	RG	JT	SC	LH	P		
health	211	57	53	210	240	207	207	60	156	
fullness	62	36	62	-45	50	112	0	0	35	
fun	-10	-22	6	12	-24	-12	3	-26	-9	
social	10	31	32	-15	-13	4	17	34	12	
energy	38	96	47	81	80	-39	153	-5	56	

Table 2: Effects of setting each basic need status to zero on percentage change in time spent on activities fulfilling those needs, relative to normal settings. *Initials of characters* - JL: John Lin, EL: Eddy Lin, MG: Monica Gellor, RG: Rachel Greene; JT: Joey Tribbiani, SC: Sheldon Cooper, LH: Leonard Hofstadter, P: Penny

As shown in Table 2, humanoid agents adapt their activities most when the basic needs of health (156%), energy (56%) and fullness (35%) are initialized to zero. This supports their importance, as Maslow (1943) grouped them into low-level physiological and safety needs that people need to satisfy before fulfilling other needs. In response to these conditions, agents typically look for medical support, get more rest or seek more food. On the other hand, when agents feel lonely due to the lack of social interactions, they only slightly adapt their behavior (+12%) to interact more with other agents.

Another contributory factor to the small changes for both social and fun is that agents in normal settings already spend a large amount of time doing activities that contribute to these basic needs: on average, they spend 11 (out of 18 simulated hours) doing something they enjoy, 8.75 hours on social interactions and only 5.75 hours for resting and 2.75 hours each for eating and doing something that improves their health.³ This means that the effects of setting either fun or social to zero initially dissipates very earlier in the day, giving way to other priorities including work obligations, such as Penny working at the Cheesecake Factory.

6 Conclusion

We propose Humanoid Agents, a platform for human-like simulations of Generative Agents guided by System 1 processing including Basic Needs (*e.g.* hunger, health and energy), Emotion and Closeness in Relationships. Our platform also powers the immersive visualization of agents using our Unity WebGL Game Interface and an Interactive Analytics Dashboard.

³Activities such as meeting friends for lunch can be enjoyable, social and filling at the same time, so the total time across all basic needs can be more than 18 hours.

Limitations

Multiparty Dialogue: Our system currently only supports dialogue between two agents, even if there are more than two agents in the same location. We aim to support multi-party conversation in the future.

Synchronization of activities between agents: Activity planning is done by each agent independently and not forcibly synchronized with each other. For instance, John Lin can plan to watch a movie with Eddy Lin at 8:00pm but Eddy Lin can be calling his friends at 8:00pm. We plan to synchronize activities between agents in the future, by prompting one or more agents to update their plan if their plans are not coherent with one another.

Variability in Natural Decline of Agent Basic Needs: The rates of decrease for various basic needs is the same for all agents in our current implementation. We plan to allow these rates to be customized for each character to account for individual differences (e.g. extroverts have faster decline of social fulfillment and people who get hungry easily can have fullness reduced more quickly).

Ethics Statement

Broader impacts: Our system will help researchers such as computational social scientists to be better able to simulate human behavior in-silico before doing further studies in the real-world. This is particularly helpful if real-world studies are difficult or costly.

Risk: While our system allows simulated agents to behave more like humans, it is not perfect and should not be treated as so. Users of our simulation platform must be informed that they are working with a simulation that does not perfectly reflect human behavior in the real world, so that they do not overly trust the results of the simulation.

Acknowledgements

We thank the anonymous reviewers for their helpful comments.

References

Sieun An, Li-Jun Ji, Michael Marks, and Zhiyong Zhang. 2017. [Two sides of emotion: Exploring positivity and negativity in six basic emotions across cultures](#). *Frontiers in Psychology*, 8.

Kira S. Birditt, Laura M. Miller, Karen L. Fingerma, and Eva S. Lefkowitz. 2009. [Tensions in the parent and adult child relationship: Links to solidarity and ambivalence](#). *Psychology and Aging*, 24(2):287–295.

Antoine Bosselut, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. 2019. [COMET: Commonsense transformers for automatic knowledge graph construction](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4762–4779, Florence, Italy. Association for Computational Linguistics.

Harrison Chase. 2022. [LangChain](#).

R.I.M. Dunbar. 2009. [The social brain hypothesis and its implications for social evolution](#). *Annals of Human Biology*, 36(5):562–572.

Paul Ekman. 1992. [An argument for basic emotions](#). *Cognition and Emotion*, 6(3-4):169–200.

Barry G. Ginsberg. 1996. [Together in group therapy: Fathers and their adolescent sons](#). In *Men in groups: Insights, interventions, and psychoeducational work.*, pages 269–282. American Psychological Association.

Michael L. Hecht. 1984. [Satisfying communication and relationship labels: Intimacy and length of relationship as perceptual frames of naturalistic conversations](#). *Western Journal of Speech Communication*, 48(3):201–216.

Siddharth Karamcheti Saachi Jain Samuel Humeau Emily Dinan Tim Rocktäschel Douwe Kiela Arthur Szlam Jason Weston Jack Urbanek, Angela Fan. 2019. [Learning to speak and act in a fantasy text adventure game](#).

Liwei Jiang, Jena D. Hwang, Chandra Bhagavatula, Ronan Le Bras, Jenny Liang, Jesse Dodge, Keisuke Sakaguchi, Maxwell Forbes, Jon Borhardt, Saadia Gabriel, Yulia Tsvetkov, Oren Etzioni, Maarten Sap, Regina Rini, and Yejin Choi. 2022. [Can machines learn morality? the delphi experiment](#).

Daniel Kahneman. 2011. *Thinking, fast and slow*. Farrar, Straus and Giroux, New York.

Guohao Li, Hasan Abed Al Kader Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. 2023. [Camel: Communicative agents for "mind" exploration of large scale language model society](#).

Fangyu Liu, Emanuele Bugliarello, Edoardo Maria Ponti, Siva Reddy, Nigel Collier, and Desmond Elliott. 2021. [Visually grounded reasoning across languages and cultures](#). In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10467–10485, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

- Ruibo Liu, Ruixin Yang, Chenyan Jia, Ge Zhang, Denny Zhou, Andrew M. Dai, Diyi Yang, and Soroush Vosoughi. 2023. [Training socially aligned language models in simulated human society](#).
- A. H. Maslow. 1943. [A theory of human motivation](#). *Psychological Review*, 50:370–396.
- Rowland S. Miller. 2012. *Intimate relationships*. McGraw-Hill.
- Anton J. Nderhof. 1985. [Methods of coping with social desirability bias: A review](#). *European Journal of Social Psychology*, 15(3):263–280.
- OpenAI. 2023. [Gpt-4 technical report](#).
- OpenBMB. 2023. [AgentVerse](#).
- Joon Sung Park, Joseph C. O’Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2023. [Generative agents: Interactive simulators of human behavior](#).
- Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. 2019. [Towards empathetic open-domain conversation models: A new benchmark and dataset](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5370–5381, Florence, Italy. Association for Computational Linguistics.
- Sam G.B. Roberts and Robin I.M. Dunbar. 2011. [The costs of family and friends: an 18-month longitudinal study of relationship maintenance and decay](#). *Evolution and Human Behavior*, 32(3):186–197.
- Richard M. Ryan and Edward L. Deci. 2001. [On happiness and human potentials: A review of research on hedonic and eudaimonic well-being](#). *Annual Review of Psychology*, 52(1):141–166.
- Ashish Sharma, Adam Miner, David Atkins, and Tim Althoff. 2020. [A computational approach to understanding empathy expressed in text-based mental health support](#). In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 5263–5276, Online. Association for Computational Linguistics.
- Gravitas Significant. 2023. [Auto GPT](#).
- Guanzhi Wang, Yuqi Xie, Yunfan Jiang, Ajay Mandlekar, Chaowei Xiao, Yuke Zhu, Linxi Fan, and Anima Anandkumar. 2023. [Voyager: An open-ended embodied agent with large language models](#).
- Zhilin Wang and Pablo E. Torres. 2022. [How to be helpful on online support forums?](#) In *Proceedings of the 4th Workshop of Narrative Understanding (WNU2022)*, pages 20–28, Seattle, United States. Association for Computational Linguistics.
- Zhilin Wang, Xuhui Zhou, Rik Koncel-Kedziorski, Alex Marin, and Fei Xia. 2022. [Extracting and inferring personal attributes from dialogue](#). In *Proceedings of the 4th Workshop on NLP for Conversational AI*, pages 58–69, Dublin, Ireland. Association for Computational Linguistics.
- Xiaodong Wu, Weizhe Lin, Zhilin Wang, and Elena Rastorgueva. 2020. [Author2vec: A framework for generating user embedding](#).
- Nakajima Yohei. 2022. [BabyAGI](#).
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. [Personalizing dialogue agents: I have a dog, do you have pets too?](#)
- W.-X. Zhou, D. Sornette, R. A. Hill, and R. I. M. Dunbar. 2005. [Discrete hierarchical organization of social group sizes](#). *Proceedings of the Royal Society B: Biological Sciences*, 272(1561):439–444.

A Appendices

A.1 Effects of Emotions on Activities

To understand the effects of emotions on the activity of agents, we apply a similar approach as we do with basic needs in setting the initial emotion to one other than neutral. Unlike basic needs however, emotion does not contain gradation and therefore changes much more quickly than basic needs statuses. An agent who wakes up sad can become neutral right after washing up, severely limiting the effects of setting the initial emotion on activities over the rest of the day. To overcome this issue, we disable the updating of emotions based on activities and dialogues for agents in this part of the study, making the agents resemble humans who are stuck in a particular emotion over the entire day.

We study the number of times in one simulated day to which agents perform activities (at 15-min intervals) that express each emotion. For instance, when agents are *angry*, they go for a run to release anger; when *sad*, they seek support from a trusted friend; when *disgusted*, they practice deep breathing exercises and meditation techniques; and when *surprised*, they take time to process and reflect on the surprising findings. Then, we calculate the difference to the number of times that agents perform such activities in normal settings (initialized with neutral and allowed to update based on activities and dialogues). Because agents in normal settings typically do not perform activities that express sadness, anger, fear, disgust or surprise, we report the increase in number of activities expressing those emotions, compared to agents in normal settings.

As shown in Table 3, anger influences agent behavior most (+15 activities) followed by sadness

Emotion	change in no. of activities expressing emotion								
	Lin's Family		Friends			Big Bang Theory			Mean
	JL	EL	MG	RG	JT	SC	LH	P	μ
angry	7	15	18	18	12	22	23	8	15
sad	6	22	8	8	15	9	8	6	<u>10</u>
afraid	3	16	4	1	12	5	14	24	<u>10</u>
disgusted	5	1	16	0	2	4	0	5	4
surprised	0	1	0	0	0	3	4	2	1
happy	6	0	-9	-9	-18	-4	-6	16	-3

Table 3: Effects of fixing agent emotion on the change in number of activities (in 15-minute intervals) that agents perform expressing set emotion, relative to normal settings. *Initials of characters* - JL: John Lin, EL: Eddy Lin, MG: Monica Gellor, RG: Rachel Greene; JT: Joey Tribbiani, SC: Sheldon Cooper, LH: Leonard Hofstadter, P: Penny

and fear (+10 each), then disgust (+4) and surprise (+1) and finally happiness (-2). Negative emotions seem to influence agents much more than positive emotions (*i.e.* happiness), as agents often do not plan for activities with negative emotions and hence have to significantly adjust their plans to manage their negative emotions, as with converging evidence in humans (An et al., 2017). Among the negative emotions, disgust and surprise are transient emotions that humans and agents do not typically experience for a long time, therefore limiting their influence compared to persistent ones (sadness, fear and anger) that can affect their activity over the entire day. It is interesting to observe that agents do slightly fewer (-2) activities to make themselves happy when they are already happy, possibly because feeling happy empowers them to pursue activities serving other longer term goals, even if these activities are not immediately joy-inducing (Ryan and Deci, 2001). For instance, when emotion is set at happy, Joey Tribbiani devotes more time to practicing his craft as an actor and Sheldon Cooper spends more time doing research.

A.2 Effects of Closeness on Dialogues

We investigate the effects of initial relationship closeness on dialogue between two agents. We set the closeness between all pairs of characters to either 0 (distant), 5 (rather close), 10 (close) or 15 (very close). We consider only the first five conversations between all agents in each simulation, in order to ensure that the closeness values between two agents are kept at the initial set level. We report the effects on the mean number of turns in each conversation as well as the proportion of turns that our system rates to have positive sentiments, since closeness between agents can influence their

Closeness	Mean turns			% positive sentiment		
	LF	F	BBT	LF	F	BBT
Distant	5.4	6.8	8.33	88.9	100	100
Rather Close	7.25	7.8	8.8	93.1	97.4	100
Close	6.2	8.2	9.8	96.8	97.5	98.0
Very Close	6.0	7.2	8.0	90.0	97.2	97.5

Table 4: Effects of closeness between agents on the mean number of turns in their conversations and the proportion of turns with positive sentiments. *Initials to World* - LF: Lin's Family, F: Friends, BBT: Big Bang Theory

willingness to express positive/negative sentiment in their conversations with one another.

As shown in Table 4, the mean number of conversation turns typically follow an inverse U shape as closeness increases. Agents talk less when they are distant, more when they get closer, but then tapers off when they get very close. This is supported by converging evidence in human conversations where we feel less of a need to engage in politeness talk when we feel very close to others (Hecht, 1984). In Lin's Family (LF), the turning point is at Rather Close while in Friends (F) and Big Bang Theory (BBT), it is at Close. This is likely between the two agents in LF have a father-son relationship where agents feel comfortable communicating less at a lower level of closeness without straining the relationship (Ginsberg, 1996) while in F and BBT, agents are friends and neighbours who require more active communication to maintain as non-kin relationships (Roberts and Dunbar, 2011).

Proportion of conversational turns with positive sentiment generally goes down in Table 4 when closeness is higher, akin to how people feel less of a need to praise others in order to be liked when very close to others (Miller, 2012). For instance, when Joey feels distant from Monica, he says 'Hey Monica, I saw you having a blast at that restaurant with your friends! The food looked incredible. What's the secret to finding such amazing places to eat? I could use some recommendations for my next date night.' When he feels very close, he instead says 'Hey Monica! I saw you having lunch at that nearby restaurant. How was the food? I'm always on the lookout for new culinary trends to try out in my cooking. Any recommendations or standout dishes you enjoyed?'

In LF, the proportion of turns with positive senti-

ment is also lower (88.9%) when agents feel distant from each other. This is possibly because when a father-child pair feel distant from each other, they can be more likely to argue with each other out of dissatisfaction with the strained relationship (Birditt et al., 2009). For instance, after dinner, Eddy Lin says to John Lin ‘Hey John, I noticed you had someone else clean up the dinner dishes tonight. Everything okay? Is there a reason for that?’ when he feels distant but says ‘I noticed how much you enjoyed that cup of tea while we were talking and bonding, and it made me appreciate our connection even more.’ when he feels very close. On the other hand, F/BBT agents entirely refrain from demonstrating negative sentiments when they are distant from their acquaintances/neighbours to maintain a positive social image, similar to humans (Nederhof, 1985). Overall, this suggests that the effect of relationship closeness on agents’ conversations is substantially moderated by their relationship type, as supported by converging evidence from human conversations.

A.3 Templates for Evaluating Basic Needs, Emotion, Closeness and Sentiment

Basic Needs Does the activity {activity} involve {satisfaction-action}? Please respond only with either yes or no.

where satisfaction-action for each basic need is based on Table 5

Emotion In the following activity {activity}, what emotion is expressed? Please respond only with one word from this list [“neutral”, “disgusted”, “afraid”, “sad”, “surprised”, “happy”, “angry”].

Closeness Given this conversation {conversation}, did {name} enjoy the conversation? Please respond with either yes or no.

Sentiment In the following utterance {utterance}, is the sentiment positive? Please respond only with either yes or no.

Basic Need	Satisfaction-action
fullness	eating food
social	interacting with other people
fun	doing something enjoyable
health	doing something that improves their own physical health
energy	resting or having a break

Table 5: Satisfaction-action for each basic need