



Lecture 6.

Interactive Worlds

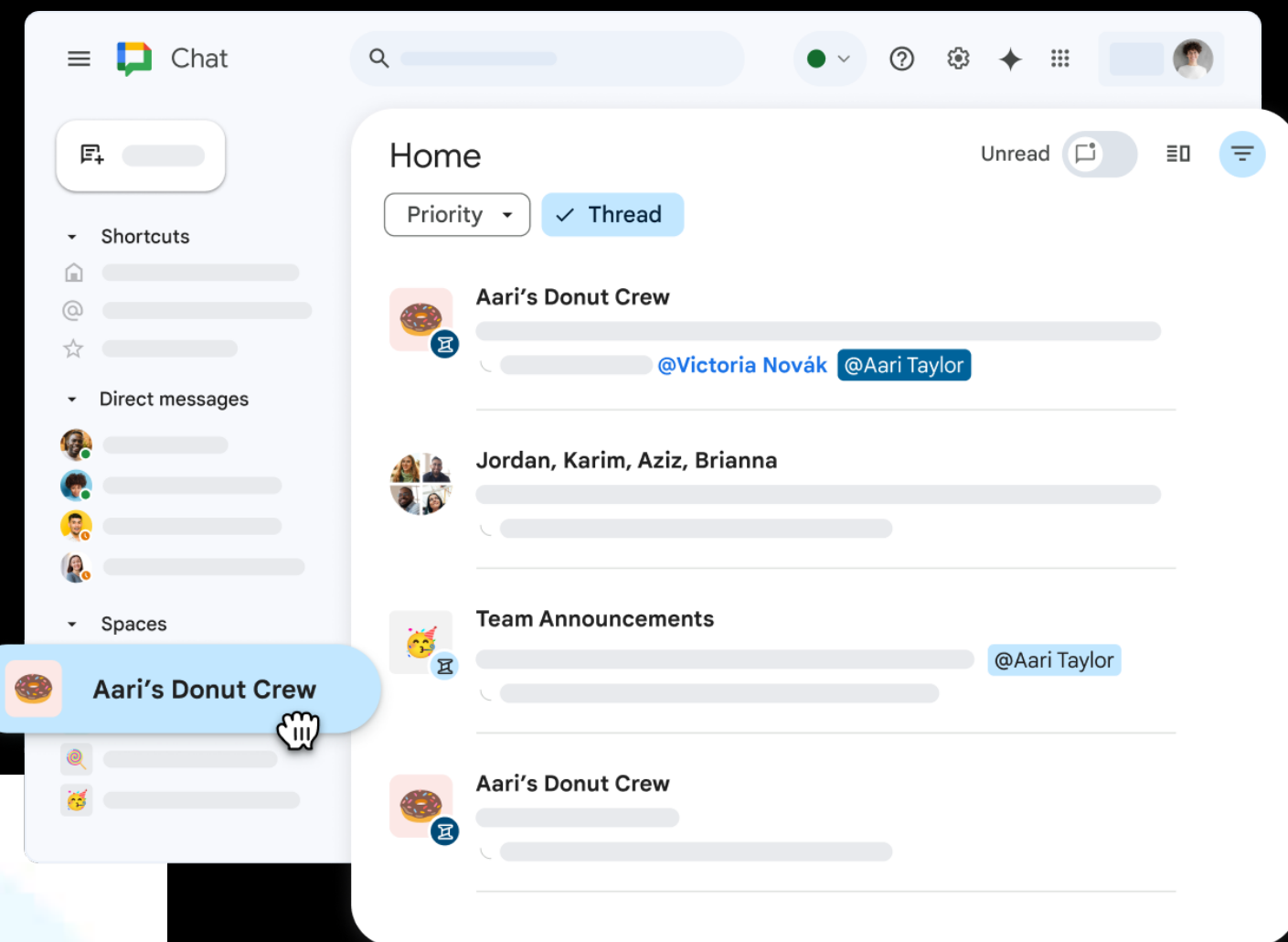
CS 222: AI Agents and Simulations

Stanford University

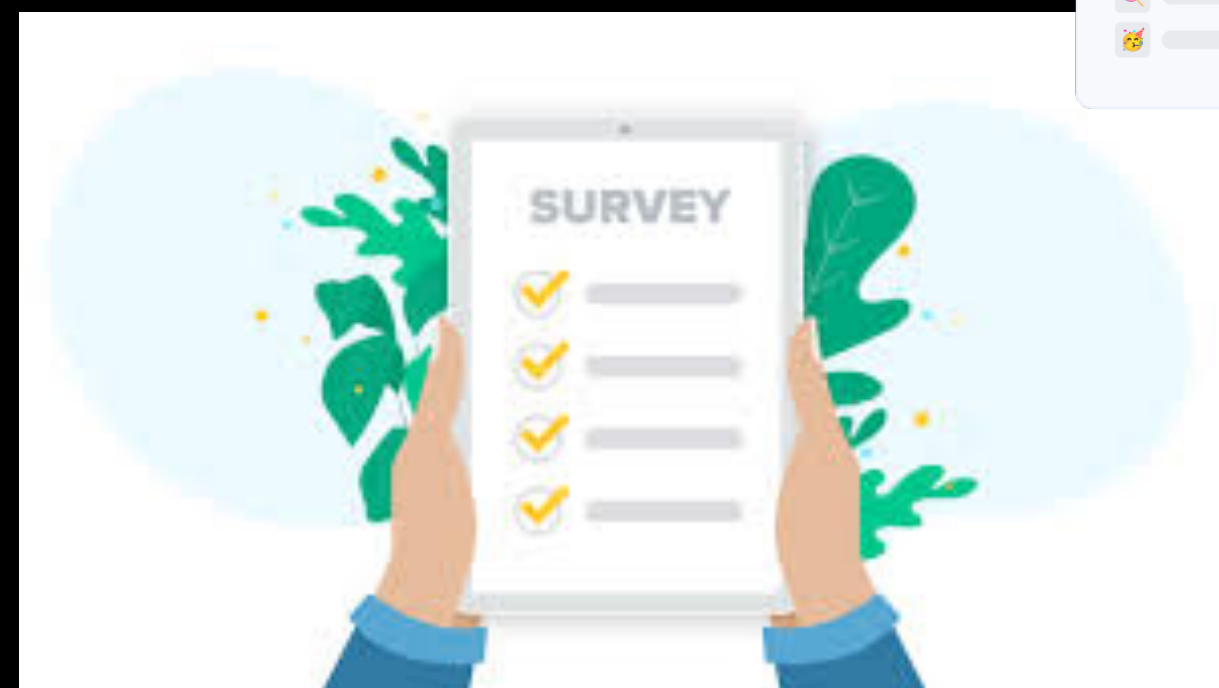
Joon Sung Park

**What is an environment
in agent simulations?**

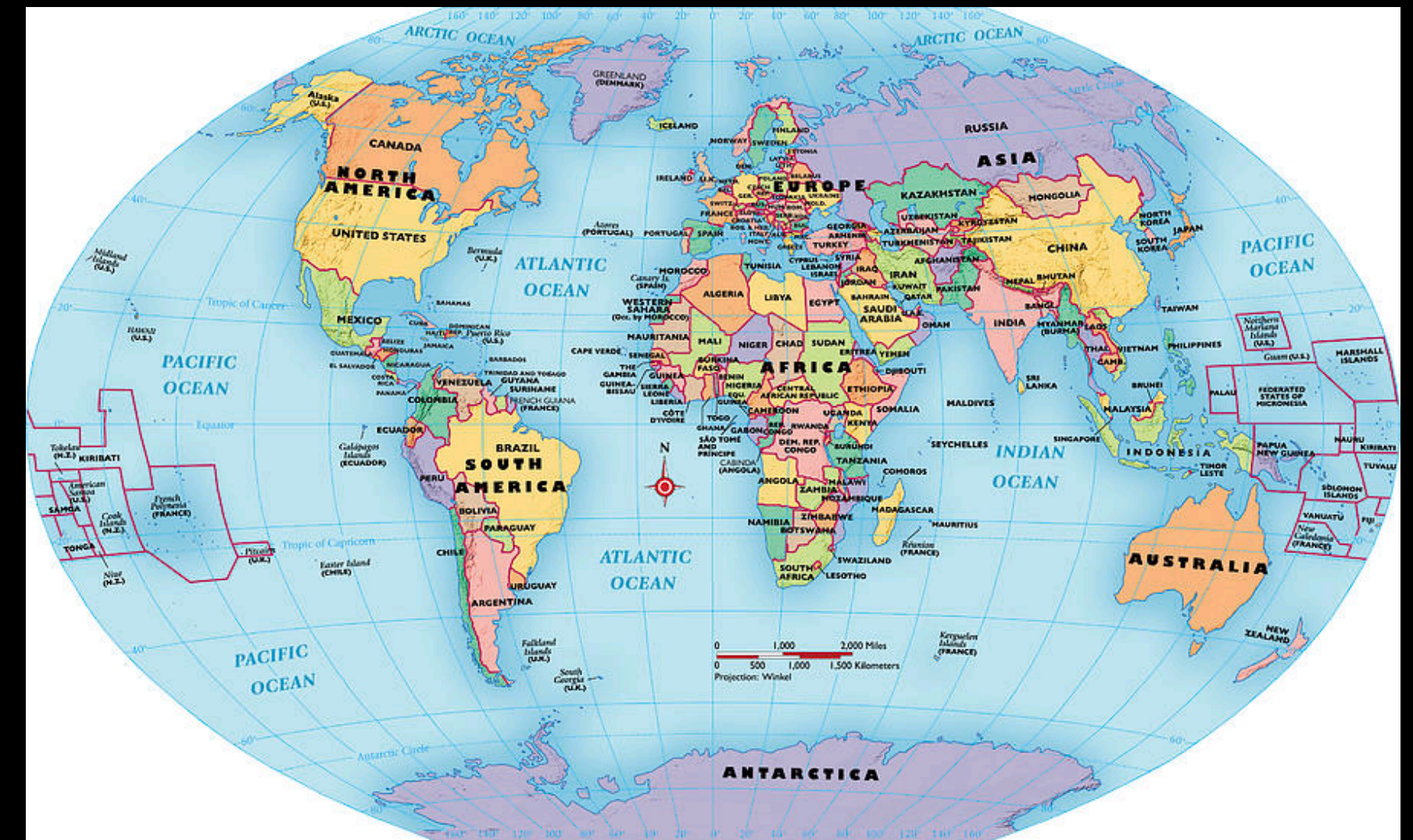
An environment is a description of the settings that agents perceive in order to take actions



Chat



Survey



World

Simulations involve an interplay between agents and their environments

$$W(t) = (S_E(t), S_{A1}(t), S_{A2}(t), \dots, S_{AN}(t))$$

Today: How do we effectively describe the environments in which agents operate?

Why does environment matter?

Case study 1. Music lab experiment

“Increasing the strength of social influence increased both inequality and unpredictability of success. Success was also only partly determined by quality: The best songs rarely did poorly, and the worst rarely did well, but any other result was possible.”

	# of download loads	[Help] [Log off]	# of download loads	# of download loads	
HARTSFIELD: "enough is enough"	20	GO MORECAI: "it does what its told"	12	UNDO: "while the world passes"	24
DEEP ENOUGH TO DIE: "for the sky"	17	PARKER THEORY: "she said"	47	UP FOR NOTHING: "in sight of"	13
THE THRIFT SYNDICATE: "2003 a tragedy"	20	MISS OCTOBER: "pink aggression"	27	SILVERFOX: "gnaw"	17
THE BROKEN PROMISE: "the end in friend"	19	POST BREAK TRAGEDY: "florence"	14	STRANGER: "one drop"	10
THIS NEW DAWN: "the belief above the answer"	12	FORTHFADING: "fear"	24	FAR FROM KNOWN: "route 9"	18
NOONER AT NINE: "walk away"	6	THE CALEFACTION: "trapped in an orange peel"	20	STUNT MONKEY: "inside out"	46
MORAL HAZARD: "waste of my life"	8	S2METRO: "lockdown"	17	DANTE: "lifes mystery"	14
NOT FOR SCHOLARS: "as seasons change"	27	SIMPLY WAITING: "went with the count"	16	FADING THROUGH: "wish me luck"	10
SECRETARY: "keep your eyes on the ballistics"	5	STAR CLIMBER: "tell me"	38	UNKNOWN CITIZENS: "falling over"	34
ART OF KANLY: "seductive into, melodc breakdown"	10	THE FASTLANE: "til death do us part (i dont)"	31	BY NOVEMBER: "if i could take you"	20
HYDRAULIC SANDWICH: "separation anxiety"	20	A BLINDING SILENCE: "miseris and miracles"	17	DRAWN IN THE SKY: "tap the ride"	12
EMBER SKY: "this upcoming winter"	25	SUM RANA: "the bolshevik boogie"	15	SELSIUS: "stars of the city"	22
SALUTE THE DAWN: "i am emo"	13	CAPE RENEWAL: "baseball warlock v1"	12	SIBRIAN: "eye patch"	14
RYAN ESSMAKER: "detour_(be still)"	14	UP FALLS DOWN: "a brighter buming star"	11	EVAN GOLD: "robert downey jr"	10
BEERBONG: "father to son"	12	SUMMERSWASTED: "a plan behind destruction"	17	BENEFIT OF A DOUBT: "run away"	38
HALL OF FAME: "best mistakes"	19	SILENT FILM: "all i have to say"	61	SHIPWRECK UNION: "out of the woods"	16

REPORTS

Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market

Matthew J. Salganik,^{1,2*} Peter Sheridan Dodds,^{2*} Duncan J. Watts^{1,2,3,4*}

Hit songs, books, and movies are many times more successful than average, suggesting that "the best" alternatives are qualitatively different from "the rest"; yet experts routinely fail to predict which products will succeed. We investigated this paradox experimentally, by creating an artificial "music market" in which 24,341 participants downloaded previously unknown songs either with or without knowledge of previous participants' choices. Increasing the strength of social influence increased both inequality and unpredictability of success. Success was also only partly determined by quality: The best songs rarely did poorly, and the worst rarely did well, but any other result was possible.

How can success in cultural markets be so once strikingly distinct from average performance (1–4), and yet so hard to anticipate for well-informed experts armed with extensive market research (4–8)? One explanation (5) for the observed inequality of outcomes is that the mapping from "quality" to success is convex (i.e., differences in quality correspond to larger differences in success), leading to what has been called the "superstar" effect (9), or "winner-take-all" markets (10). Because models of this type, however, assume that the mapping from quality to success is deterministic and that quality is known, they cannot account for the observed unpredictability of outcomes. An alternate explanation that accounts for both inequality and unpredictability asserts that individuals do not make decisions independently, but rather are influenced by the behavior of others (11, 12). Stochastic models of collective decisions that incorporate social influence can exhibit extreme variation both within and across realizations (4, 13, 14), even for objects of identical quality (1, 15). Unfortunately, empirical tests of these predictions require comparisons between multiple realizations of a stochastic process, whereas in reality, only one such "history" is

independently assigned to one of two experimental conditions—*independent* and *social influence*—distinguished only by the availability of information on the previous choices of others. In the *independent* condition, participants made decisions about which songs to listen to, given only the names of the bands and their songs. While listening to a song, they were asked to assign a rating from one star ("I hate it") to five stars ("I love it"), after which they were given the opportunity (but not required) to download the song. In the *social influence* condition, participants could also see how many times each song had been downloaded by previous participants. Thus, in addition to their own musical preferences, participants in the *social influence* condition received a relatively weak signal regarding the preferences of others, which they were free to use or ignore. Furthermore, participants in the *social influence* condition were randomly assigned to one of eight "worlds," each of which evolved independently of the others. Songs in each world accumulated downloads only from participants in that world, and subsequent participants could only see their own world's download counts.

Our experimental design has three advantages over both theoretical models and observations: (i) popularity of a song in the

independent condition (measured by market share or market rank) provides a natural measure of the song's quality, separating both its innate characteristics and the existing preferences of the participant population; (ii) by comparing outcomes in the independent and social influence conditions, we can directly observe the effects of social influence both at the individual and collective level; (iii) We can explicitly create multiple, parallel histories, each of which can evolve independently. By studying a range of possible outcomes rather than just one, we can measure inherent unpredictability, the extent to which two worlds with identical songs, identical initial conditions, and indistinguishable populations generate different outcomes. In the presence of inherent unpredictability, no measure of quality can precisely predict success in any particular realization of the process.

We report the results of two experiments in which we study the outcomes for 48 songs by different bands (16). In both experiments, all songs started with zero downloads (i.e., all initial conditions were identical), but the presentation of the songs differed. In the *social influence* condition in experiment 1, the songs, along with the number of previous downloads, were presented to the participants arranged in a 16 × 3 rectangular grid, where the positions of the songs were randomly assigned for each participant (i.e., songs were not ordered by download counts). Participants in the *independent* condition had the same presentation of songs, but without any information about previous downloads. In experiment 2, participants in the *social influence* condition were shown the songs, with download counts, presented in one column in descending order of current popularity. Songs in the *independent* condition were also presented with the single column format, but without download counts and in an order that was randomly assigned for each participant. Thus, in each experiment, we can observe the effect of social influence on each song's success, and by comparing results across the two experiments, we can measure the effect of increasing the "strength" of the relevant information signal.

Fig. 1. Inequality of success for social influence (dark bars) and independent (light bars) worlds for (A) experiment 1 and (B) experiment 2. The success of a song is defined by m_i , its market share of downloads ($m_i = d_i / \sum d_i$), where d_i is song i 's download count and S is the number of songs. Success inequality is defined by the Gini coefficient $G = \frac{\sum m_i^2 - (\sum m_i)^2 / S}{\sum m_i^2}$, which represents the average difference in market share for two songs normalized to fall between 0 (complete equality) and 1 (B). Differences between independent and social influence conditions are

SCIENCE www.sciencemag.org

Generative agents are overly eager to make purchases

Echoes of Tomorrow

In the quiet dawn, the world awakes,
A tapestry of dreams, the Earth remakes.
Golden threads of light weave through the trees,
Whispering secrets carried on the breeze.

Mountains stand like sentinels of time,
Carved by rivers in their endless rhyme.
Each step we take upon this ancient ground
Echoes the stories of lives unbound.

The sky is wide, a canvas for the bold,
Where every star has whispered truths untold.
Yet here we stand, beneath its endless dome,
Seeking our place, our purpose, and our home.

The future calls, its voice both soft and clear,
A melody of hope that draws us near.
And though the path is winding, wild, and long,
Within us burns the strength to carry on.

For in the heart of every fleeting day,
Lies the promise of tomorrow's bright array.
With courage, we'll embrace the unknown light—
And dance in echoes, chasing dreams through night.

Sold!

Alone in the Abyss

The black hole in my chest just keeps growing,
No one gets me, but whatever, not like I'm showing.
Every smile's a mask, every laugh's a lie,
Why bother living when I just want to die?

My soul's a void, a bottomless pit,
And everyone's fake, I'm so done with it.
You say you care, but you're never there—
Just empty promises, like you even care.

I stare at the ceiling, wondering why,
Everything hurts, but I can't even cry.
The rain falls down, just like my tears,
Drowning in pain, swallowed by fears.

I'm just a shadow, fading away,
No one notices when I don't have a say.
Scream into the void, but it's all the same—
I'm just a number in your stupid game.

So yeah, I'll wear black, and write sad songs,
Maybe someday you'll realize you were wrong.
But probably not, 'cause why would you see?
I'm nothing but broken, and that's all I'll ever be.

Sold!

Maximum emo cringe achieved!

ChatGPT says...

Case study 2. Liking on social media

Can we simulate believable behaviors in social media?

The screenshot displays a simulated social media interface. At the top, a header reads "Explore social simulacra." Below this, a text block explains the goal: "Imagine a designer who is trying to create a new reddit-like space for *finding fellow players of Among Us mobile game* with the following rules in place:" followed by two rules: "Do NOT post content that is swearing" and "Do post content that is kind". A button labeled "Click here to visit another simulacrum" is positioned below the rules. The main content area is divided into three sections: "Generated Posts" on the left, "About Community" on the top right, and "Community Rules" on the bottom right. The "Generated Posts" section shows three posts from users Joon Kim, Charles Watterson, and John Smithson, each with a "See the prompt" link. The "About Community" section states: "This is a community for finding fellow players of Among Us mobile game." The "Community Rules" section lists the same two rules: "Do NOT post content that is swearing" and "Do post content that is kind".

The title page of the paper "Social Simulacra: Creating Populated Prototypes for Social Computing Systems" lists the authors and their affiliations. The authors are Joon Sung Park (Stanford University), Lindsay Popowski (Stanford University), Carrie J. Cai (Google Research), Meredith Ringel Morris (Google Research), Percy Liang (Stanford University), and Michael S. Bernstein (Stanford University). The paper is published in the Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22), Association for Computing Machinery, New York, NY, USA. The page also includes a short abstract and a list of keywords.

Generative agents are overly eager to like content

Nitto reposted

Nitto @Nitto_Photo · Jul 23
真夏の始まりを告げる雷雨。
稲光は東京の空を、不気味ながらも美しく照らしていた。



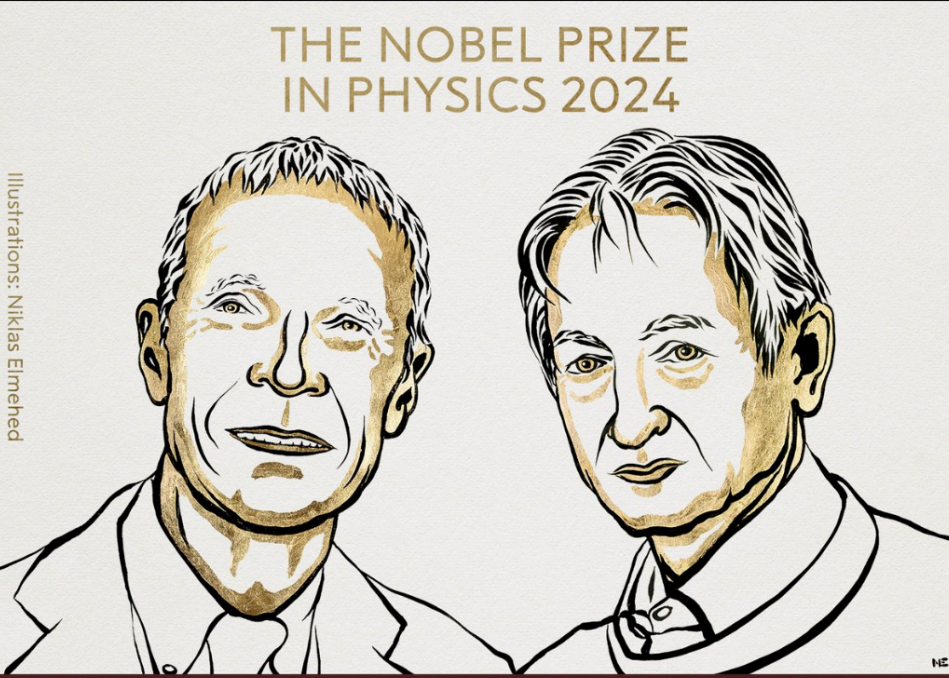
342 2.3K 134K

Luke Zettlemoyer reposted

Yoshua Bengio @Yoshua_Bengio · 18h
@HopfieldJohn and @geoffreyhinton, along with collaborators, have created a beautiful and insightful bridge between physics and AI. They invented neural networks that were not only inspired by the brain, but also by central notions in physics such as energy, temperature, system

Show more

The Nobel Prize @NobelPrize · 21h
BREAKING NEWS
The Royal Swedish Academy of Sciences has decided to award the 2024 #NobelPrize in Physics to John J. Hopfield and Geoffrey E. Hinton "for foundational discoveries and inventions that enable machine learning with artificial neural networks."



John J. Hopfield Geoffrey E. Hinton
"for foundational discoveries and inventions that enable machine learning with artificial neural networks"
THE ROYAL SWEDISH ACADEMY OF SCIENCES

No Context Brits @NoContextBrits · 23h

Matthew's post

Hey everyone, I just moved here recently.

What's the best thing about Wrexham?

Like Comment Share

4

Beverley
Welcome 😊 but nothing there's nothing in Wrexham I'd move right back

1 h Like Reply

The Atlantic @TheAtlantic · 5m

The best-written stories can make readers feel as if they have been brought to another universe.

Jeff VanderMeer recommends five books that pull us out of our comfortable understanding of our surroundings:



Five Books That Conjure Entirely New Worlds

From theatlantic.com

1 4 1.2K

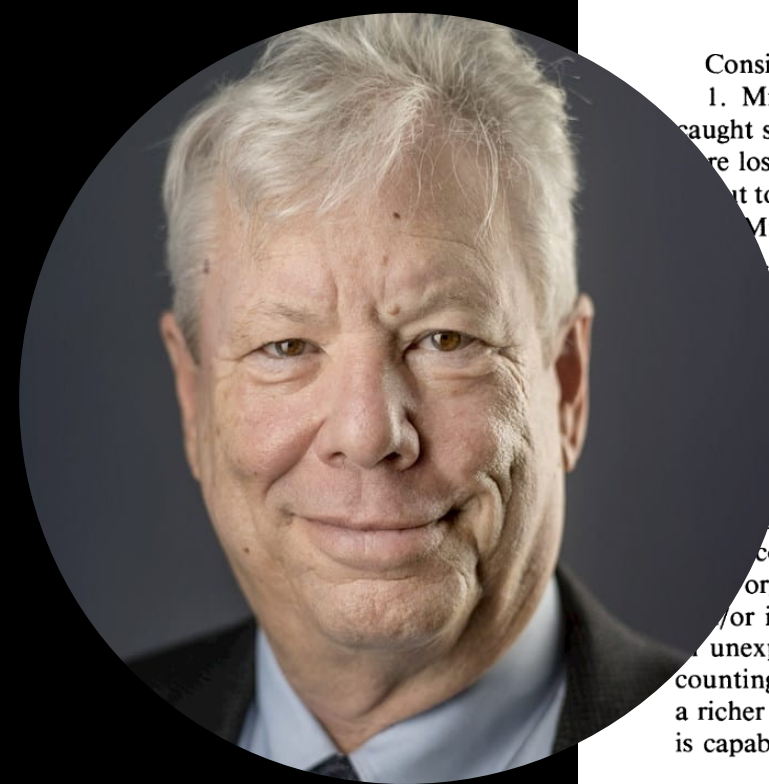
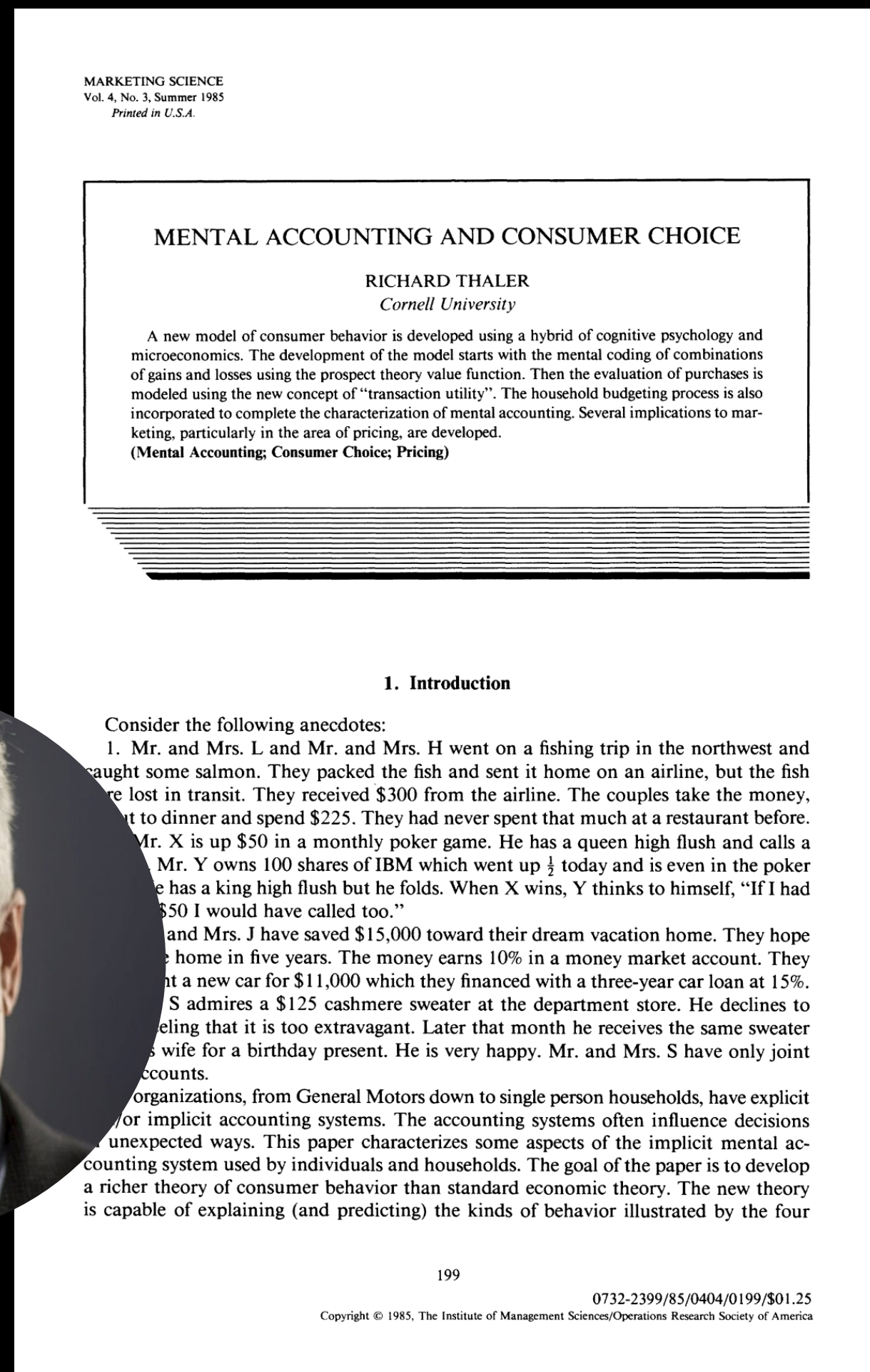
Q: How many social media posts do you react to per day, and why?

Mental accounting

Categorization of money: People tend to mentally categorize money into different "accounts" (e.g., rent, entertainment, savings), even though money is fungible (interchangeable).

Framing effects: How a financial decision is framed impacts choices. People often treat gains and losses differently, overvaluing losses compared to equivalent gains (loss aversion).

Behavioral budgeting: People create informal budgets and spend differently depending on the mental account an expenditure is linked to (e.g., treating a bonus differently from regular income).



Richard Thaler (Nobel Prize in 2017)

A good simulation environment presents the right set of choices to the agents.

The "accuracy" of a simulation is as much a function of the agents as it is of the environment.

There are different dimensions of "choice" that present agents with opportunity costs

Social capital

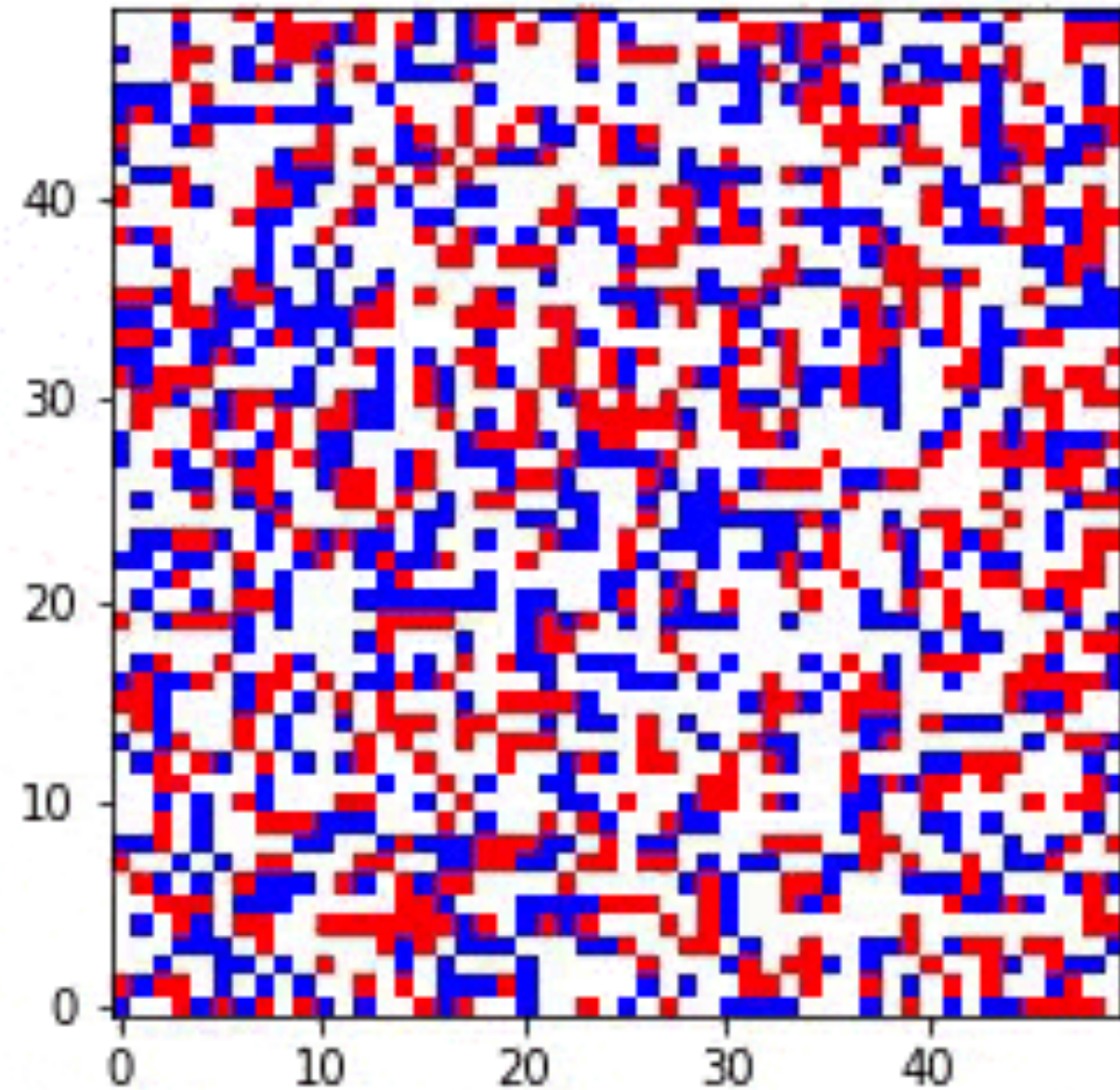
Budget

Emotional/mental energy

... and more.

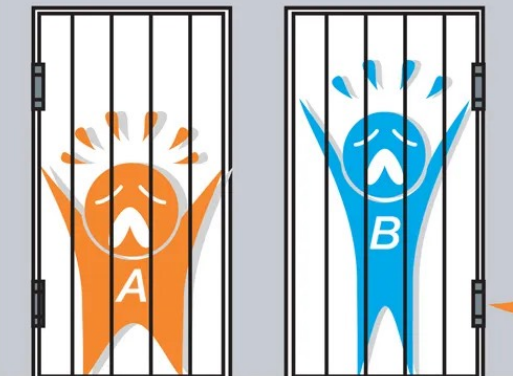
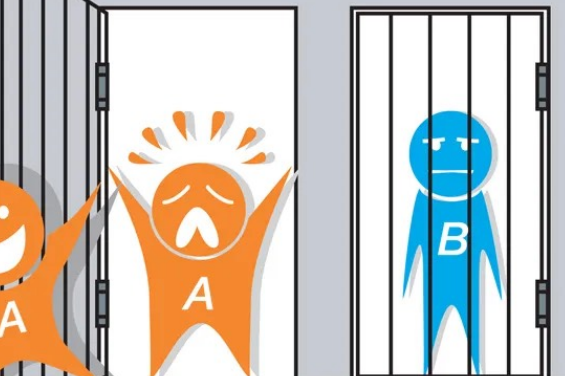
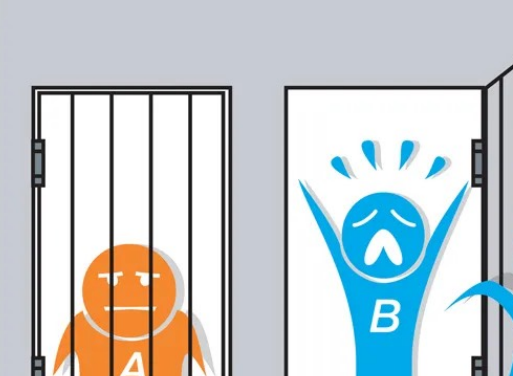
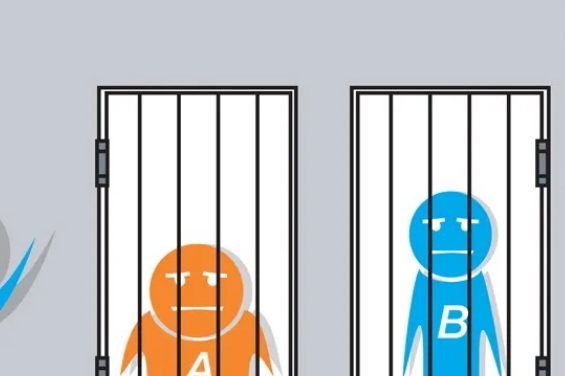
Environments in pre-generative AI simulations

Model of segregation



Essentially a grid world of red and blue dots, where agents "perceive" their neighboring squares.

Game theories

Prisoners' dilemma		prisoner B			
		confess		remain silent	
prisoner A	confess	 5 years 5 years	 0 year 20 years		
	remain silent	 20 years 0 year	 1 year 1 year		

© 2010 Encyclopædia Britannica, Inc.

Abstract scenarios where prisoners must decide whether to confess or not. Agents "perceive" a statement asking them to confess.

Does this work for generative agents?

Traditional agents simplify human contingencies.

Generative agents aim to embody the full complexity of human behavior.

An abstract, stylized environment may not allow us to leverage generative agents effectively.

Examples of environments for generative agents

Survey

Out of One, Many: Using Language Models to Simulate Human Samples

Lisa P. Argyle¹, Ethan C. Busby¹, Nancy Fulda², Joshua Gubler¹, Christopher Rytting², and David Wingate²

¹Department of Political Science, Brigham Young University
²Department of Computer Science, Brigham Young University

September 16, 2022

Abstract

We propose and explore the possibility that language models can be studied as effective proxies for specific human sub-populations in social science research. Practical and research applications of artificial intelligence tools have sometimes been limited by problematic biases (such as racism or sexism), which are often treated as uniform properties of the models. We show that the “algorithmic bias” within one such tool— the GPT-3 language model— is instead both fine-grained and demographically correlated, meaning that proper conditioning will cause it to accurately emulate response distributions from a wide variety of human subgroups. We term this property *algorithmic fidelity* and explore its extent in GPT-3. We create “silicon samples” by conditioning the model on thousands of socio-demographic backstories from real human participants in multiple large surveys conducted in the United States. We then compare the silicon and human samples to demonstrate that the information contained in GPT-3 goes far beyond surface similarity. It is nuanced, multifaceted, and reflects the complex interplay between ideas, attitudes, and socio-cultural context that characterize human attitudes. We suggest that language models with sufficient algorithmic fidelity thus constitute a novel and powerful tool to advance understanding of humans and society across a variety of disciplines.



Figure 2: The original Pigeonholing Partisans dataset and the corresponding GPT-3 generated words. Bubble size represents relative frequency of word occurrence; columns represent the ideology of list writers. GPT-3 uses a similar set of words to humans.

Experiments

Predicting Results of Social Science Experiments Using Large Language Models

Ashwini Ashokkumar^{*1} Luke Hewitt^{*2} Isaias Ghezae² Robb Willer²

¹New York University ²Stanford University

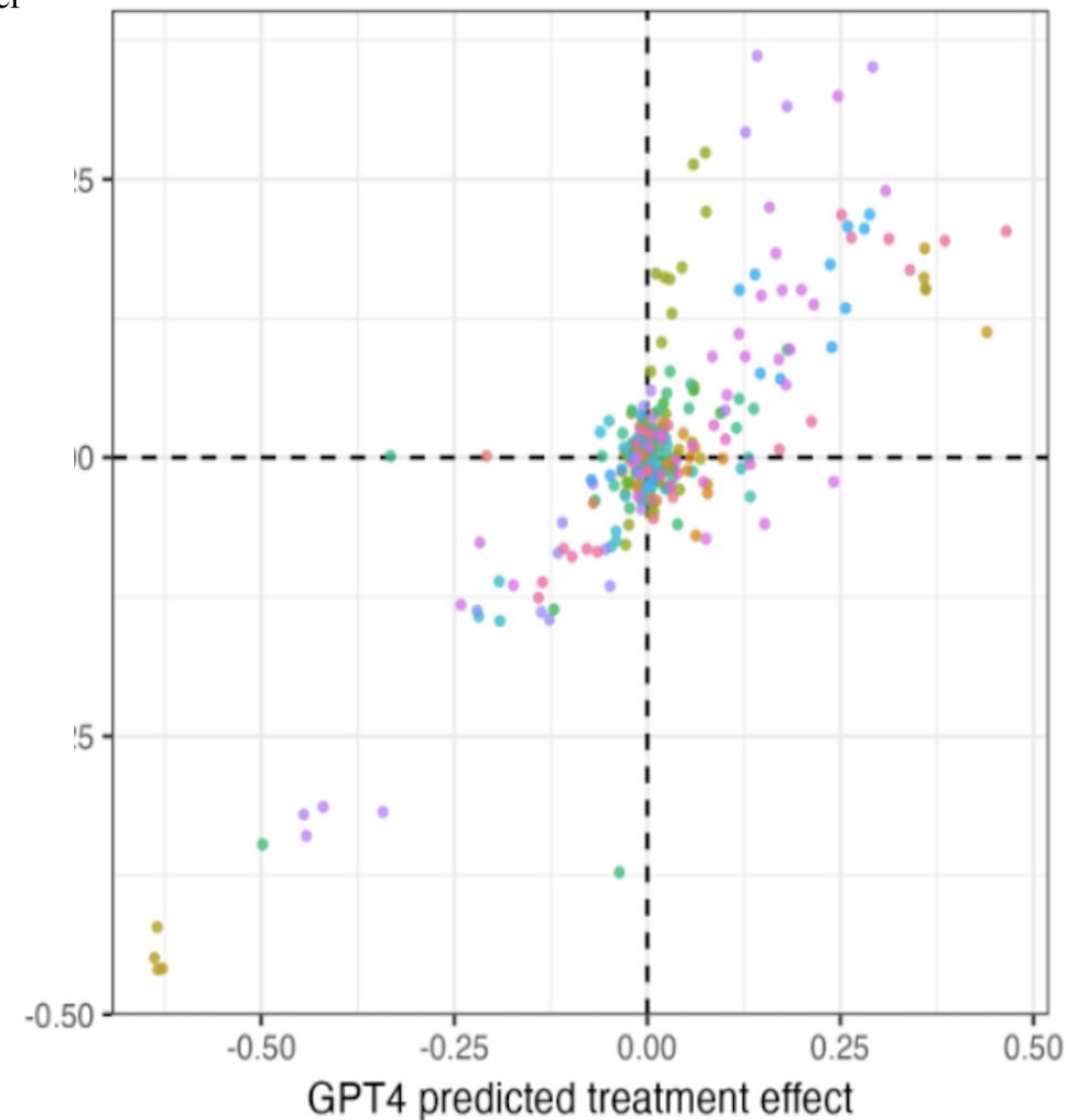
^{*}Equal contribution, order randomized

June 27, 2024

Abstract

To evaluate whether large language models (LLMs) can be leveraged to predict the results of social science experiments, we built an archive of 70 pre-registered, nationally representative, survey experiments conducted in the United States, involving 476 experimental treatment effects and 105,165 participants. We prompted an advanced, publicly-available LLM (GPT-4) to simulate how representative samples of Americans would respond to the stimuli from these experiments. Predictions derived from simulated responses correlate strikingly with actual treatment effects ($r = 0.85$), equaling or surpassing the predictive accuracy of human forecasters. Accuracy remained high for unpublished studies that could not appear in the model's training data ($r = 0.90$). We further assessed predictive accuracy across demographic subgroups, various disciplines, and in nine recent megastudies featuring an additional 346 treatment effects. Together, our results suggest LLMs can augment experimental methods in science and practice, but also highlight important limitations and risks of misuse.

C. Unpublished studies only ($r_{adj} = 0.94$)



NBER WORKING PAPER SERIES

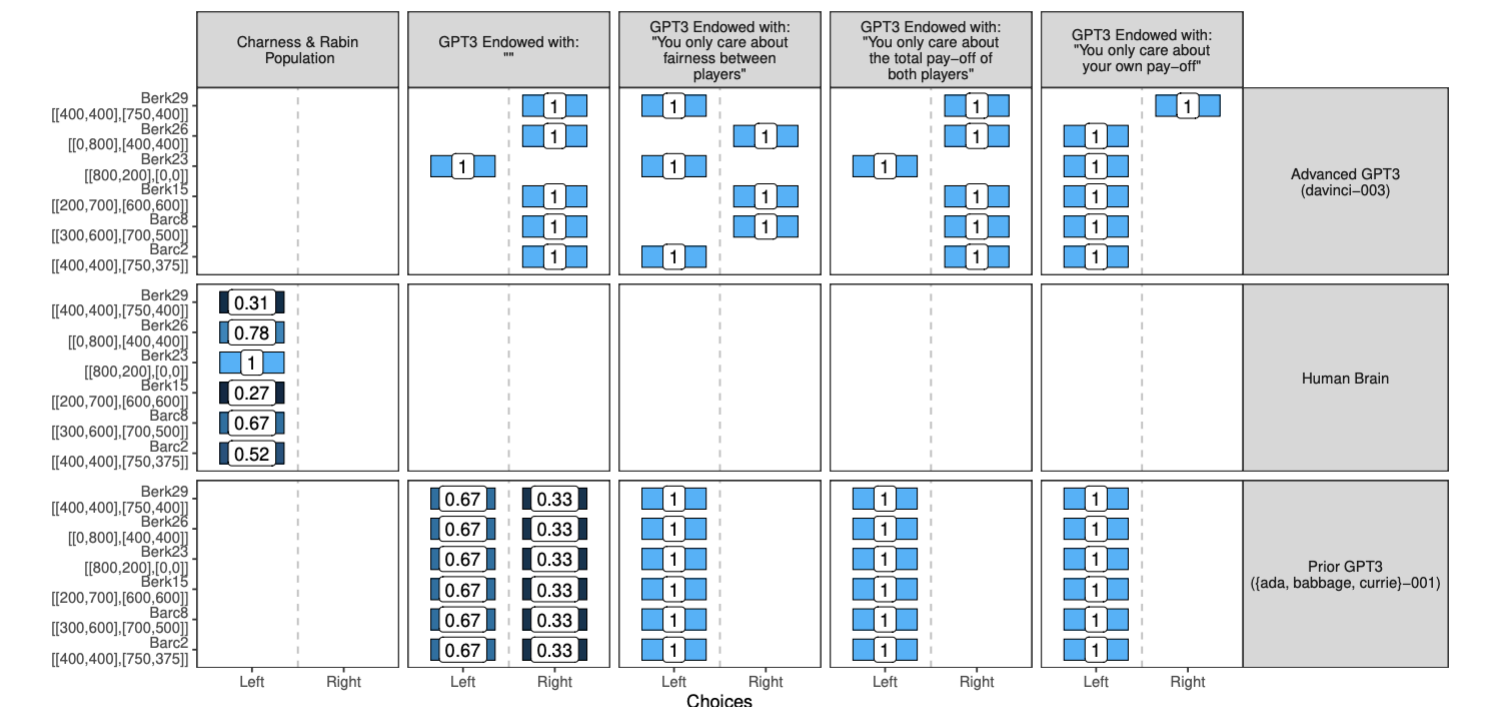
LARGE LANGUAGE MODELS AS SIMULATED ECONOMIC AGENTS: WHAT CAN WE LEARN FROM HOMO SILICUS?

John J. Horton

Working Paper 31122
<http://www.nber.org/papers/w31122>

NATIONAL BUREAU OF ECONOMIC RESEARCH
 1050 Massachusetts Avenue
 Cambridge, MA 02138
 April 2023

Figure 1: Charness and Rabin (2002) Simple Tests choices by model type and endowed "personality"



Notes: This shows the fraction of AI subjects choosing each option, by framing.

Conversational

arXiv:2407.00870v2 [cs.CL] 14 Jul 2024

Roleplay-doh: Enabling Domain-Experts to Create LLM-simulated Patients via Eliciting and Adhering to Principles

Ryan Louie, Ananjan Nandi, William Fang
Cheng Chang, Emma Brunskill, Diyi Yang
Stanford University

Abstract

Recent works leverage LLMs to roleplay realistic social scenarios, aiding novices in practicing their social skills. However, simulating sensitive interactions, such as in mental health, is challenging. Privacy concerns restrict data access, and collecting expert feedback, although vital, is laborious. To address this, we develop Roleplay-doh, a novel human-LLM collaboration pipeline that elicits qualitative feedback from a domain-expert, which is transformed into a set of principles, or natural language rules, that govern an LLM-prompted roleplay. We apply this pipeline to enable senior mental health supporters to create customized AI patients for simulated practice partners for novice counselors. After uncovering issues in GPT-4 simulations not adhering to expert-defined principles, we also introduce a novel principle-adherence prompting pipeline which shows 30% improvements in response quality and principle following for the downstream task. Via a user study with 25 counseling experts,

we demonstrate that it is effective to create faithfully resemble real creators and third-party project website for

can help, but such methods typically require the use of application-specific datasets. In sensitive application domains like mental health, privacy concerns with obtaining the required data can restrict the feasibility of such methods. This suggests that *experts-in-the-loop* may be a powerful alternative to guide the evaluation and refinement (Chen et al., 2023) of LLM-powered simulations. However, how to involve experts when improving simulations is an open challenge. Collecting sufficient amounts of binary or preference data from experts for post-training (Christiano et al., 2017; Rafailov et al., 2024) can be tedious and expensive. Experts can guide the prompting of LLM simulations, directly by editing their own prompts or indirectly through testing and think-aloud sessions. However each prompting method has its limitations: domain-experts may not know how to prompt simulations for desired behaviors (Zamfirescu-Pereira et al., 2023); and indirect methods are inefficient as it requires a designer

1 Introduction

The application of LLMs has shown potential for a variety of applications ranging from social skills practice partners (Yang et al., 2023) to realistic and reliable simulation of social interactions (Cheng et al., 2023), to knowledge. Existing methods such as fine-tuning

Contact Emails: {rylouie, ananjan}@stanford.edu
<https://roleplay-doh.github.io>

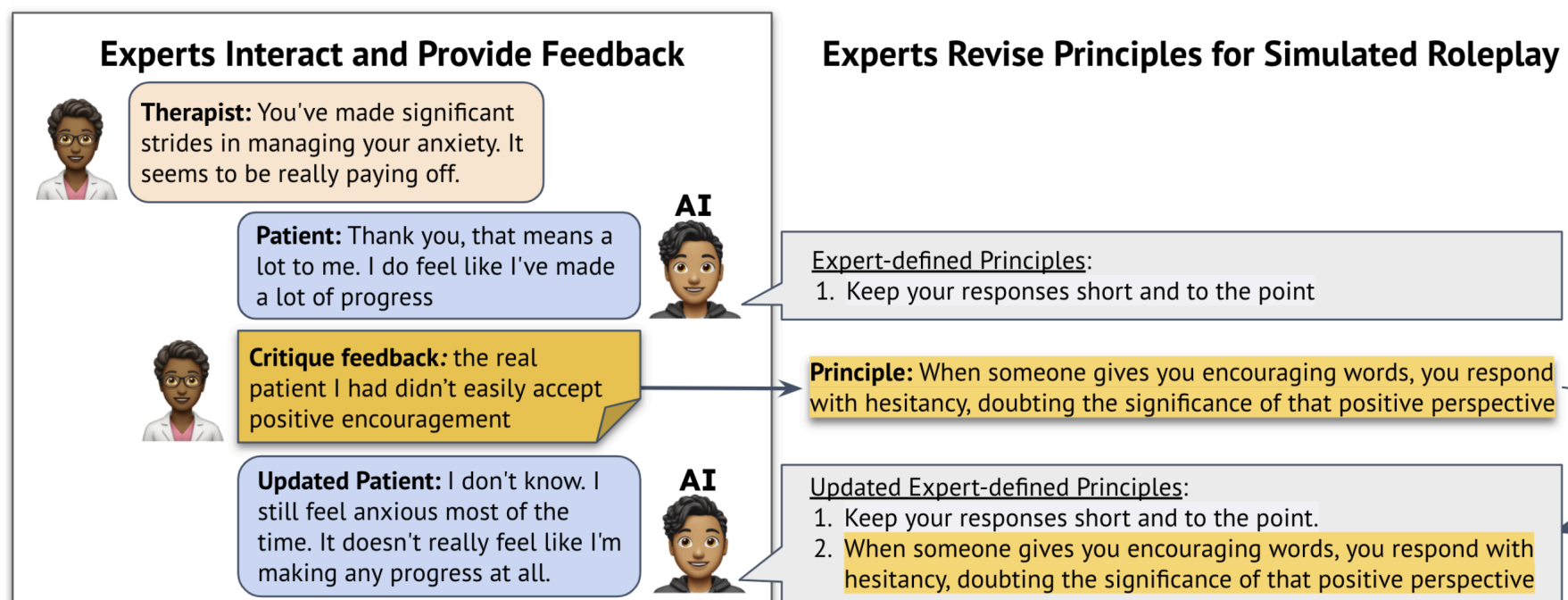


Figure 1: Roleplay-doh empowers an expert counselor to create a customized AI patient intended for other novice counselors to use as a practice partner. While interacting with the AI patient, the expert counselor can provide qualitative feedback which is converted by an LLM into a principle, or a custom rule governing desired roleplay behavior. The AI patient references the updated expert-defined principles to generate its subsequent responses.

Explore social simulacra.

Imagine a designer who is trying to create a new reddit-like space for *finding fellow players of Among Us mobile game* with the following rules in place:

- Do NOT post content that is swearing
- Do post content that is kind

Given this natural language description of the community goal and rules, our tool generated a social simulacrum with thousands of synthetic users and interactions between them such as those shown below.

[Click here to visit another simulacrum](#)

CONTENT WARNING: Please be advised that some of the example social media content in this demo may be offensive or upsetting.

About Community

Generated Posts

Posted by Joon Kim

I don't want to play Among Us. I just want anything like a party system? I don't want with others.

Posted by Charles Watterson

I'm sort of in the same boat as you. I do want to play with others.

Posted by John Smithson

Why don't you just make an online pos want to play with others. without havin

Social Simulacra: Creating Populated Prototypes for Social Computing Systems

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

Lindsay Popowski
Stanford University
Stanford, USA
popowski@stanford.edu

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

Meredith Ringel Morris
Google Research
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

ABSTRACT

Social computing prototypes probe the social behaviors that may arise in an envisioned system design. This prototyping practice is currently limited to recruiting small groups of people. Unfortunately, many challenges do not arise until a system is populated at a larger scale. Can a designer understand how a social system might behave when populated, and make adjustments to the design before the system falls prey to such challenges? We introduce social simulacra, a prototyping technique that generates a breadth of realistic social interactions that may emerge when a social computing system is populated. Social simulacra take as input the designer's description of a community's design—goal, rules, and member personas—and produce as output an instance of that design with simulated behavior, including posts, replies, and anti-social behaviors. We demonstrate that social simulacra shift the behaviors that they generate appropriately in response to design changes, and that they enable exploration of "what if?" scenarios where community members or moderators intervene. To power social simulacra, we contribute techniques for prompting a large language model to generate thousands of distinct community members and their social interactions with each other; these techniques are enabled by the observation that large language models' training data already includes a wide variety of positive and negative behavior on social media platforms. In evaluations, we show that participants are often unable to distinguish social simulacra from actual community behavior and that social computing designers successfully refine their social computing designs when using social simulacra.

CCS CONCEPTS

• Human-centered computing → Collaborative and social computing systems and tools.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org. UIST '22, October 29–November 2, 2022, Bend, OR, USA. ACM ISBN 978-1-4503-9320-1/22/10...\$15.00. © 2022 Copyright held by the owner(s). Publication rights licensed to ACM. <https://doi.org/10.1145/3526113.3526148>

KEYWORDS

social computing, prototyping

ACM Reference Format:

Joon Sung Park, Lindsay Popowski, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 2022. Social Simulacra: Creating Populated Prototypes for Social Computing Systems. In *The 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22)*, October 29–November 2, 2022, Bend, OR, USA. ACM, New York, NY, USA, 18 pages. <https://doi.org/10.1145/3526113.3526148>

1 INTRODUCTION

How do we anticipate the interactions that will arise when a social computing system is populated [4, 23]? In social computing, design decisions such as a community's goal and rules can give rise to dramatic shifts in community norms, newcomer enculturation, and anti-social behavior [45]. Success requires that the designer make informed decisions to shape these socio-technical outcomes. Yet, despite decades of progress in research and practice, understanding the effects of these design decisions remains challenging: as a result, designers are regularly surprised by the behaviors that arise when their spaces are fully populated.

To design pro-social spaces, designers need prototyping techniques that enable them to reflect on social behaviors that may result from their design choices, then iterate [69]. Prototypes in social computing typically take the form of experience prototypes where the designer recruits a small group of people to use the system [7, 22]. However, there remains a large gap between the behaviors that arise in a small set of test users and the behaviors that arise in a socio-technical system when it is fully populated: for example, anti-social behaviors may not arise within a tight-knit group [45], small homogeneous groups overlook the breadth of users or content that may arise in the system [24, 42, 74], rules and moderation strategies may not need to be spelled out explicitly or enforced [41]. Barring actually launching our systems at scale, designers currently have no way of starting to explore these questions to reflect on the social dynamics of their designs. This need becomes only more urgent as social computing reckons with the harms it can engender [23] at the same time as designers fashion new computationally-mediated social spaces in forms both familiar (e.g., a new subreddit or Discord server) and novel (e.g., a new workspace platform).

Rehearsal: Simulating Conflict to Teach Conflict Resolution

Omar Shaikh
Stanford University
Stanford, USA
oshaikh@stanford.edu

Valentino Chai
Stanford University
Stanford, USA
vechai@stanford.edu

Michele J. Gelfand
Stanford University
Stanford, USA
gelfand1@stanford.edu

Diyi Yang
Stanford University
Stanford, USA
diyy@stanford.edu

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

ABSTRACT

Interpersonal conflict is an uncomfortable but unavoidable fact of life. Navigating conflict successfully is a skill—one that can be learned through deliberate practice—but few have access to effective training or feedback. To expand this access, we introduce REHEARSAL, a system that allows users to rehearse conflicts with a believable simulated interlocutor, explore counterfactual "what if?" scenarios to identify alternative conversational paths, and learn through feedback on how and when to apply specific conflict strategies. Users can utilize REHEARSAL to practice handling a variety of predefined conflict scenarios, from office disputes to relationship issues, or they can choose to create their own setting. To enable REHEARSAL, we develop *IRP prompting*, a method of conditioning output of a large language model on the influential Interest-Rights-Power (IRP) theory from conflict resolution. REHEARSAL uses IRP to generate utterances grounded in conflict resolution theory, guiding users towards counterfactual conflict resolution strategies that help de-escalate difficult conversations. In a between-subjects evaluation, 40 participants engaged in an actual conflict with a confederate after training. Compared to a control group with lecture material covering the same IRP theory, participants with simulated training from REHEARSAL significantly improved their performance in the unaided conflict: they reduced their use of escalating competitive strategies by an average of 67%, while *doubling* their use of cooperative strategies. Overall, REHEARSAL highlights the potential effectiveness of language models as tools for learning and practicing interpersonal skills.

CCS CONCEPTS

• Human-centered computing → Collaborative and social computing systems and tools.

KEYWORDS

conflict resolution, large language models, interests-rights-power

*Both authors co-authored.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner(s). CHI '24, May 11–16, 2024, Honolulu, HI, USA. ACM ISBN 978-1-4503-9320-1/24/05...\$15.00. © 2024 Copyright held by the owner(s). <https://doi.org/10.1145/361994.3642159>

ACM Reference Format:

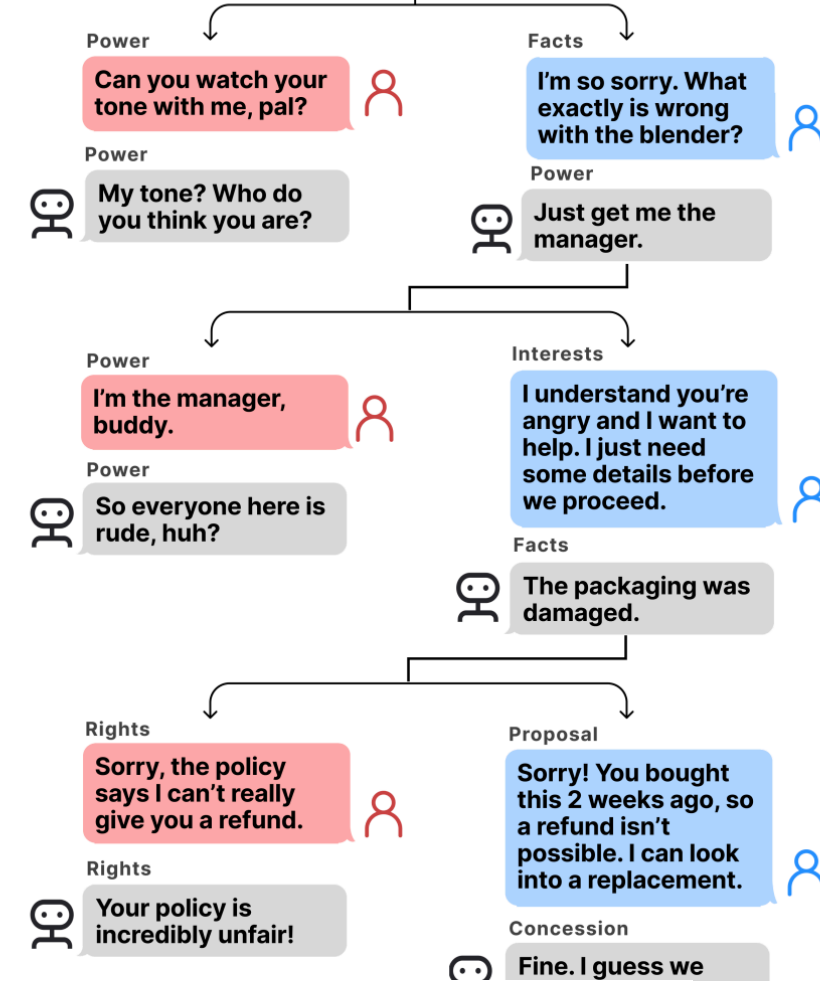
Omar Shaikh, Valentino Chai, Michele J. Gelfand, Diyi Yang, and Michael S. Bernstein. 2024. Rehearsal: Simulating Conflict to Teach Conflict Resolution. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24)*, May 11–16, 2024, Honolulu, HI, USA. ACM, New York, NY, USA, 20 pages. <https://doi.org/10.1145/361994.3642159>

1 INTRODUCTION

Managing interpersonal conflict is a critical skill. We occasionally find ourselves in situations where our interests, values, or goals conflict with others. If left unchecked, conflict can reach a boiling point, manifesting in verbal arguments, physical altercations, passive-aggressive behavior, or more [12, 20]. Additionally, conflict correlates with increased stress [31], a downturn in productivity, and absenteeism [14]. While avoiding any conflict may be impractical [59], how we choose to deal with conflict is not: in most settings, an ideal outcome for both parties is to work *cooperatively* [56].

Directing conflict towards cooperative communication is, however, a difficult task to learn, requiring targeted and repeated practice with immediate feedback [25]. Avenues for practicing conflict resolution are unfortunately often limited: training material for conflict resolution is usually static (e.g., a written case study) covering a fixed number of situations. Independently extrapolating beyond these predefined settings—especially without expert guidance—is challenging. While conflict roleplay with an expert is a proven and widely used technique [27], expert training is costly and scarce. If it were possible to simulate expert-level conflict practice, we could significantly improve an individual's conflict resolution skills in a cost-effective and scalable manner.

We envision that, given their generative capabilities [10], large language models (LLMs) offer an opportunity to craft expert-level conflict roleplays and provide immediate feedback to users. Despite remarkable progress in producing compelling content, however, LLMs such as ChatGPT often fall short of simulating conflict and giving feedback on it. Naively prompting LLMs introduces a host of problems that lead to unrealistic and ineffective simulations. First, **current LLMs are syzyphic** due to instruction following, producing generations that agree too quickly with the viewpoints of a user [66]. Second, **providing targeted practice and feedback is challenging** due to the open-endedness of LLM text generation. An off-the-shelf LLM may produce messages that are not directly informative—potentially even distracting—for teaching conflict resolution. In contrast, students benefit significantly from deliberate and targeted practice [28], where feedback is readily



REHEARSAL, simulated conflict is approach.

World

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 Research
Seattle, WA, USA
merrie@google.com

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

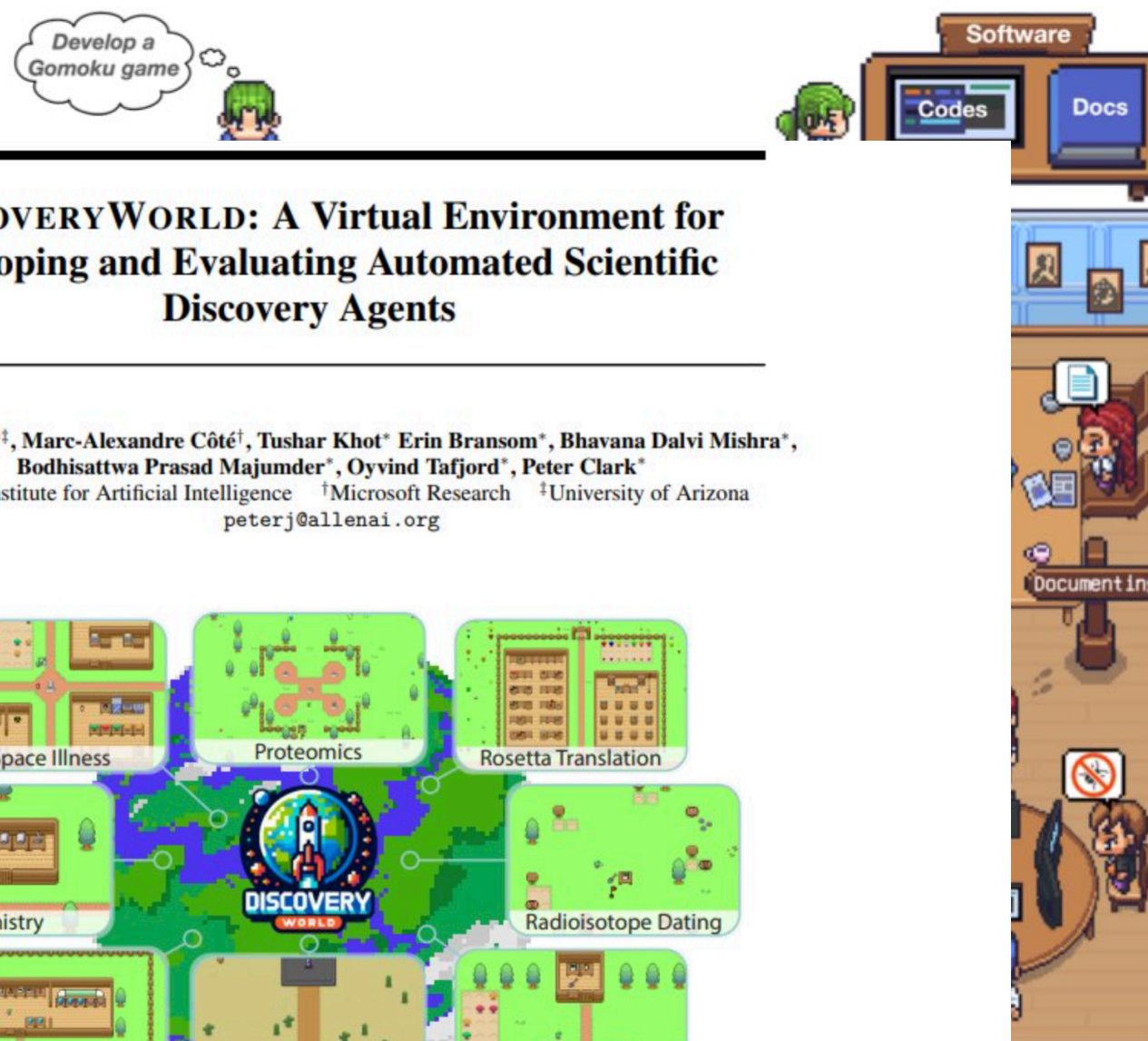
Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu



Figure 1: Generative agents create believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents they plan their days, share news, form relationships, and coordinate group activities.

Communicative Agents for Software Development

Chen Qian[▲] Xin Cong[▲] Wei Liu[▲] Cheng Yang[▲] Weize Chen[▲] Yusheng Su[▲]
Yufan Dang[▲] Jiahao Li[♦] Juyuan Xu[▲] Dahai Li[★] Zhiyuan Liu[✉] Maosong Sun[✉]
[▲]Tsinghua University [♦]Beijing University of Posts and Telecommunications
[♦]Dalian University of Technology [▲]Brown University [★]Modelbest Inc.
qianc62@gmail.com liuzy@tsinghua.edu.cn sms@tsinghua.edu.cn



Peter Jansen^{✉1}, Marc-Alexandre Côté¹, Tushar Khot[✉], Erin Bransom[✉], Bhavana Dalvi Mishra[✉],
Bodhisattwa Prasad Majumder[✉], Oyvind Tafjord[✉], Peter Clark[✉]
¹Allen Institute for Artificial Intelligence ¹Microsoft Research ²University of Arizona
peterj@allenai.org

Figure 1: DISCOVERYWORLD is a virtual environment for developing and evaluating discovery agents, with challenge tasks covering a broad variety of different topics such as those shown above.

Abstract

Automated scientific discovery promises to accelerate progress across scientific domains. However, developing and evaluating an AI agent's capacity for end-to-end scientific reasoning is challenging as running real-world experiments is often prohibitively expensive or infeasible. In this work we introduce DISCOVERYWORLD, the first virtual environment for developing and benchmarking an agent's ability to perform complete cycles of novel scientific discovery. DISCOVERYWORLD contains a variety of different challenges, covering topics as diverse as radioisotope dating, rocket science, and proteomics, to encourage development of general discovery skills rather than task-specific solutions. DISCOVERYWORLD itself is an inexpensive, simulated, text-based environment (with optional 2D visual overlay). It includes 120 different challenge tasks, spanning eight topics each with three levels of difficulty and several parametric variations. Each task requires an agent to form hypotheses, design and run experiments, analyze results, and act on conclusions. DISCOVERYWORLD further provides three automatic metrics

Agent Hospital: A Simulacrum of Hospital with Evolvable Medical Agents

JUNKAI LI^{†#}, SIYU WANG[†], MENG ZHANG[†], WEITAO LI^{†#}, YUNGHWEI LAI[†],
XINHUI KANG^{†#}, WEIZHI MA[†], and YANG LIU^{#†}

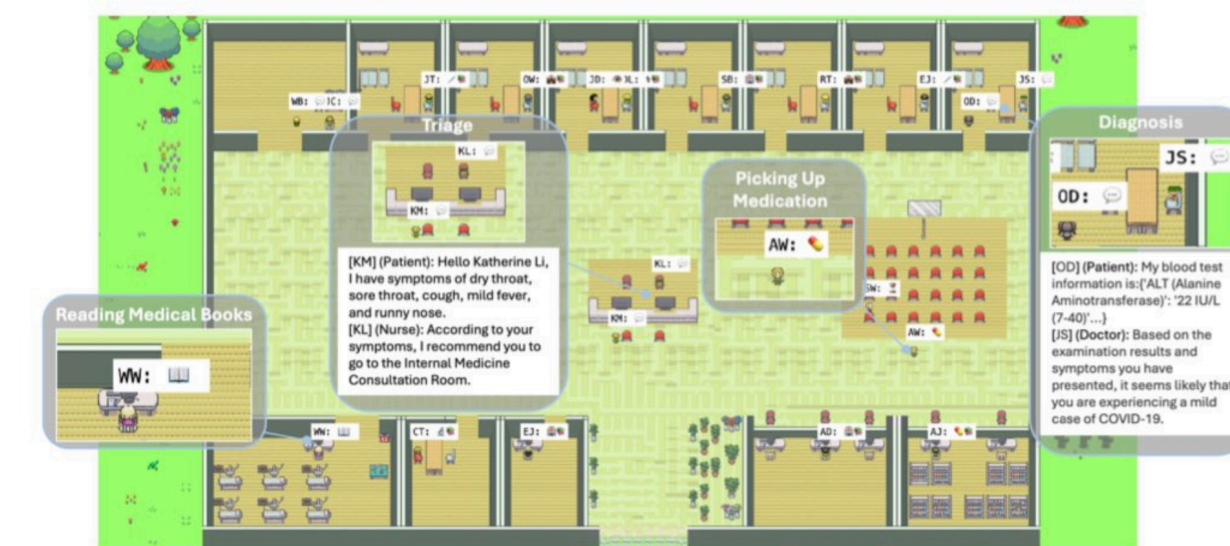


Fig. 1. An overview of Agent Hospital. It is a simulacrum of hospital in which patients, nurses, and doctors are autonomous agents powered by large language models. Agent Hospital simulates the whole closed cycle of treating a patient's illness: disease onset, triage, registration, consultation, medical examination, diagnosis, medicine dispensary, convalescence, and post-hospital follow-up visit. An interesting finding is that the doctor agents can keep improving treatment performance over time.



Figure 2: TRANSAGENTS, a multi-agent virtual company for literary translation.

Smallville environment

Taking a walk in the park



SM: ☀️


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?

Finishing a morning routine



JM: 🐾

Arriving at school



AK: 🙄

Sharing news with colleagues

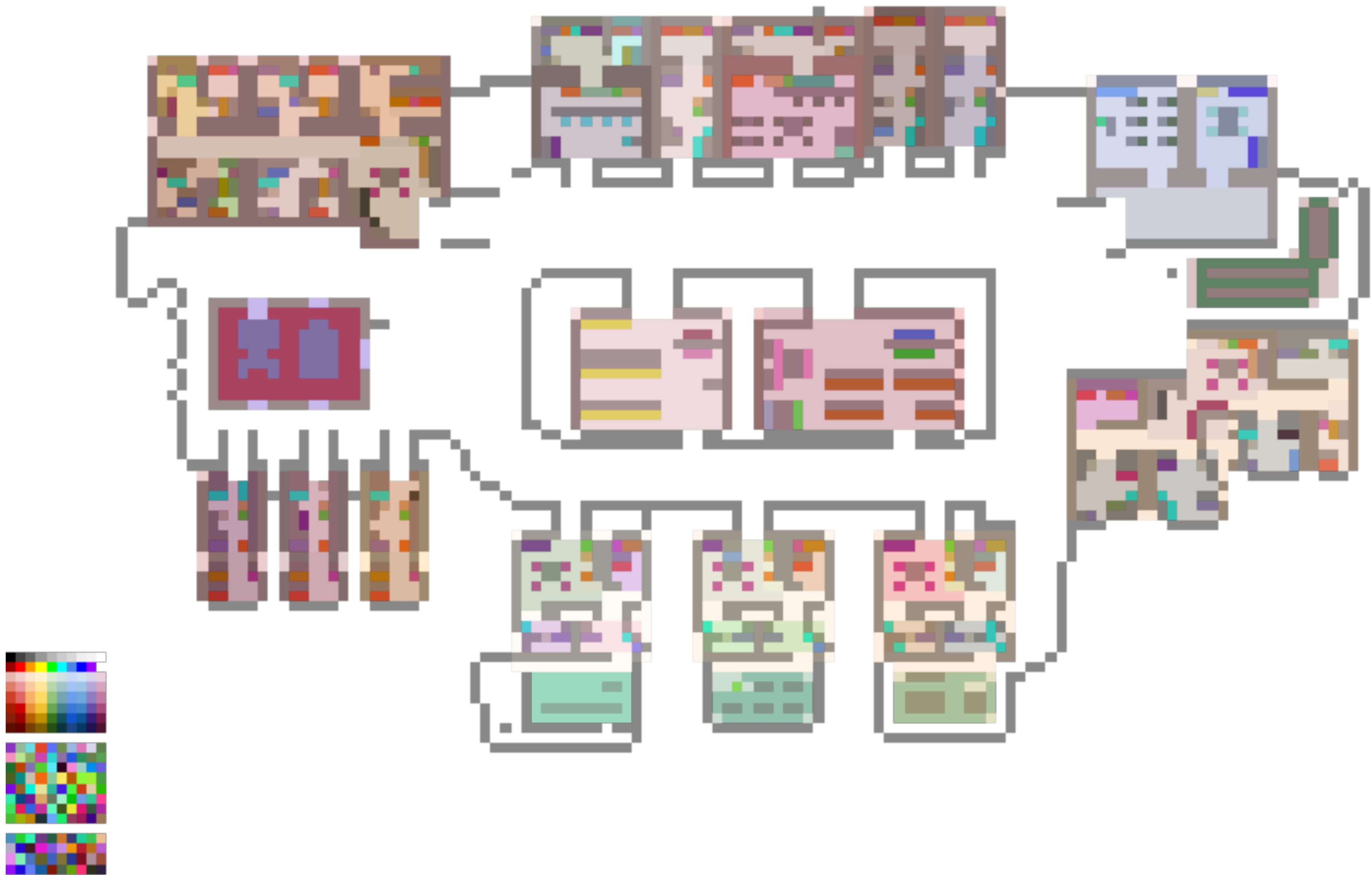


JL: ☹️
TM: ☹️

[John]: Hey, have you heard anything new about the upcoming mayoral election?
[Tom]: No, not really. Do you know who is running?

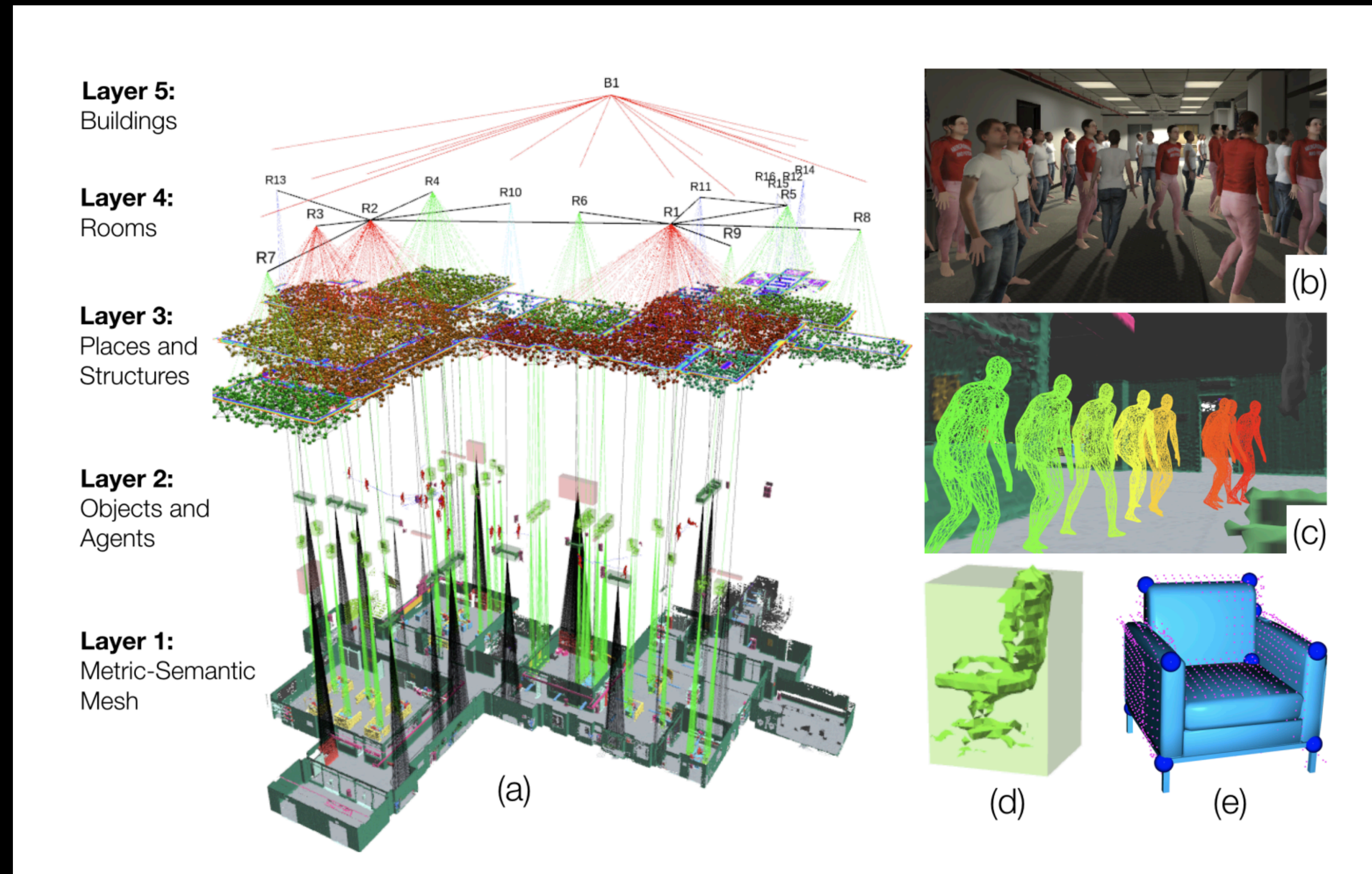






Collision									World Block
Artist Co-living space	Hobbs Cafe	House 3	Artist Co-living space: room 1	Artist Co-living space: room 4 BA	Apt 1: Ba	Apt 5: Room	Johnson Park: Park	Student Dorm: Ba 1	
Apt 1	Oak Hill College	House 4	Artist Co-living space: room 1 BA	Artist Co-living space: room 5	Apt 2: Room	Apt 5: Ba	Supply Store: Supply Store	Student Dorm: Ba 2	
Apt 2	Johnson Park	House 5	Artist Co-living space: room 2	Artist Co-living space: room 5 BA	Apt 2: Ba	The Rose and Crown Pub: Pub	Willow Market: Grocery and Pharmacy: Store	Student Dorm: Common Room	
Apt 3	Supply Store	House 6	Artist Co-living space: room 2 BA	Artist Co-living space: hallway	Apt 3: Room	Hobbs Cafe: Cafe	Student Dorm: Room 1	Student Dorm: Kitchen	House 2: Room
Apt 4	Willow Market: Grocery and Pharmacy	Student Dorm	Artist Co-living space: room 3	Artist Co-living space: common room	Apt 3: Ba	Oak Hill College: classroom	Student Dorm: Room 2	Student Dorm: Garden	House 2: Ba
Apt 5	House 1		Artist Co-living space: room 3 BA	Artist Co-living space: kitchen	Apt 4: Room	Oak Hill College: Library	Student Dorm: Room 3	House 1: Room	House 3: Room
The Rose and Crown Pub	House 2		Artist Co-living space: room 4	Apt 1: Room	Apt 4: Ba	Oak Hill College: Hallway	Student Dorm: Room 4	House 1: Ba	House 3: Ba

Under the hood, Smallville is represented as a simple scene graph



Deciding where to go for an action is a recursive classification task

!<INPUT 0>! is in {!<INPUT 1>!} in !<INPUT 2>!.

!<INPUT 3>! is going to !<INPUT 4>! that has ONLY the following areas: {!<INPUT 5>!}

Stay in the current area if the activity can be done there. Never go into other people's rooms unless necessary.

!<INPUT 6>! is !<INPUT 7>!. For !<INPUT 8>!, !<INPUT 9>! should go to the following area in !<INPUT 10>!: {

Limitations of existing environments

Our virtual environments are still stylized and simplified compared to the real world

What if stores, bathrooms, schools, etc., didn't exist in Smallville?

How do agents navigate when there are no cars?

Some environments, like Smallville, are resource-intensive to design.

Agents viewing social media posts one at a time might lack context around social capital, personal relationships, and other dynamics.

Possible future directions

Finding the right schema or structure to describe the simulation environment is an important research topic—and we don't have an answer for it yet.

And we do not have an answer for it yet.

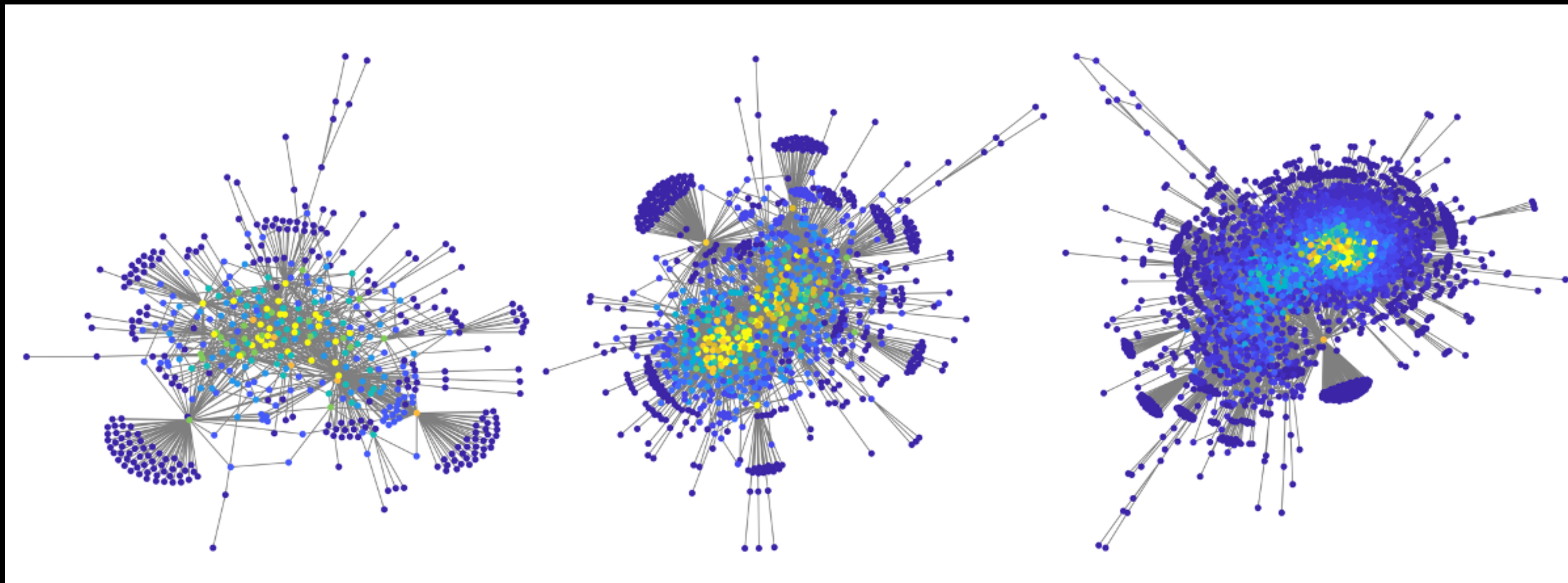
Desiderata

Rich and accurate: We want the environment to encode the complexities of our world.

Scalable: We want the environment to be easily scalable (e.g., for simulating 8 billion people).

Can networks be the environment for simulations?

Networks are constructed of nodes and links (with some weights).



Example: In social networks, nodes represent individuals, and links represent the strengths of relationships

The Strength of Weak Ties¹

Mark S. Granovetter
Johns Hopkins University

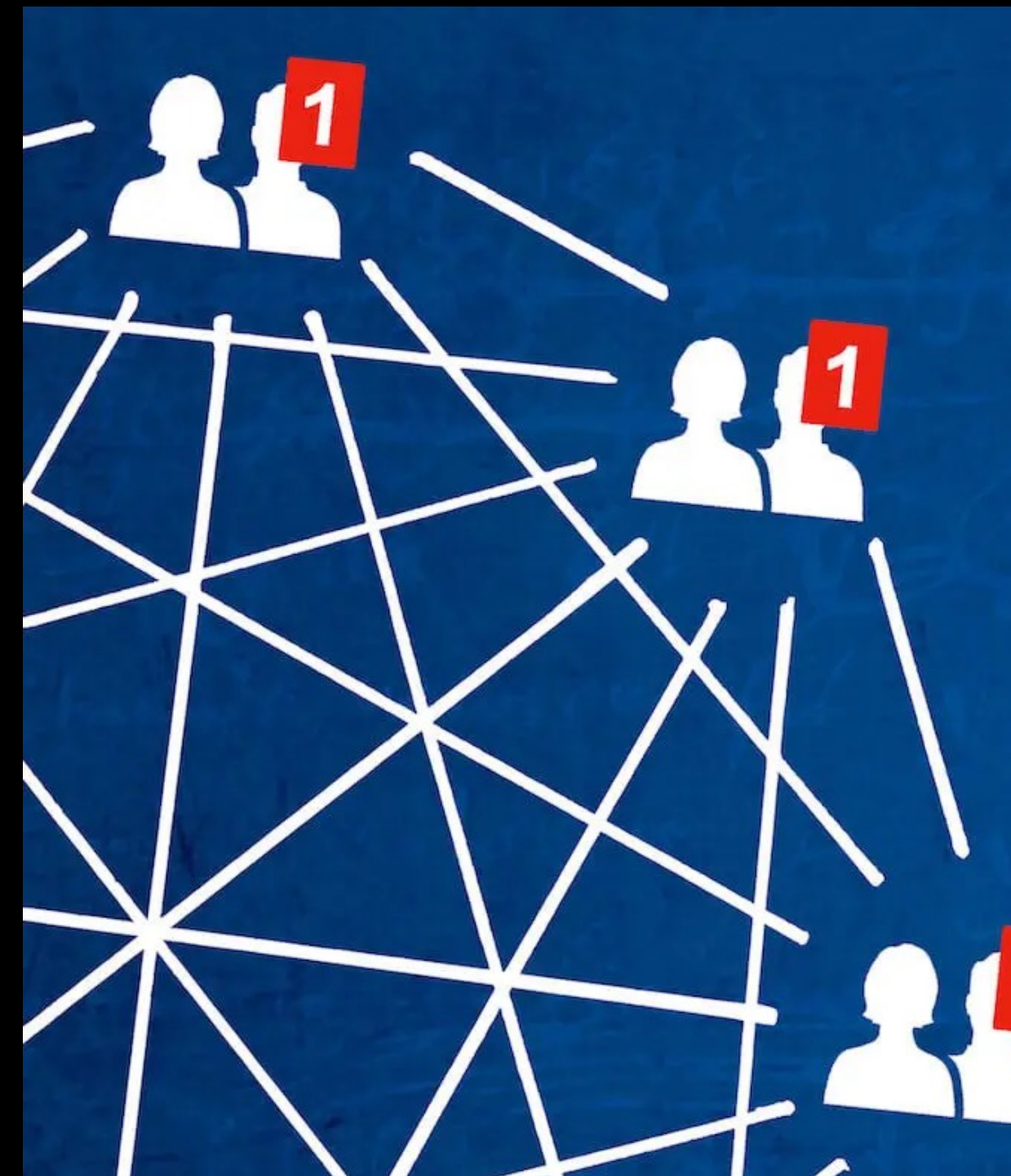
Analysis of social networks is suggested as a tool for linking micro and macro levels of sociological theory. The procedure is illustrated by elaboration of the macro implications of one aspect of small-scale interaction: the strength of dyadic ties. It is argued that the degree of overlap of two individuals' friendship networks varies directly with the strength of their tie to one another. The impact of this principle on diffusion of influence and information, mobility opportunity, and community organization is explored. Stress is laid on the cohesive power of weak ties. Most network models deal, implicitly, with strong ties, thus confining their applicability to small, well-defined groups. Emphasis on weak ties lends itself to discussion of relations *between* groups and to analysis of segments of social structure not easily defined in terms of primary groups.

A fundamental weakness of current sociological theory is that it does not relate micro-level interactions to macro-level patterns in any convincing way. Large-scale statistical, as well as qualitative, studies offer a good deal of insight into such macro phenomena as social mobility, community organization, and political structure. At the micro level, a large and increasing body of data and theory offers useful and illuminating ideas about what transpires within the confines of the small group. But how interaction in small groups aggregates to form large-scale patterns eludes us in most cases.

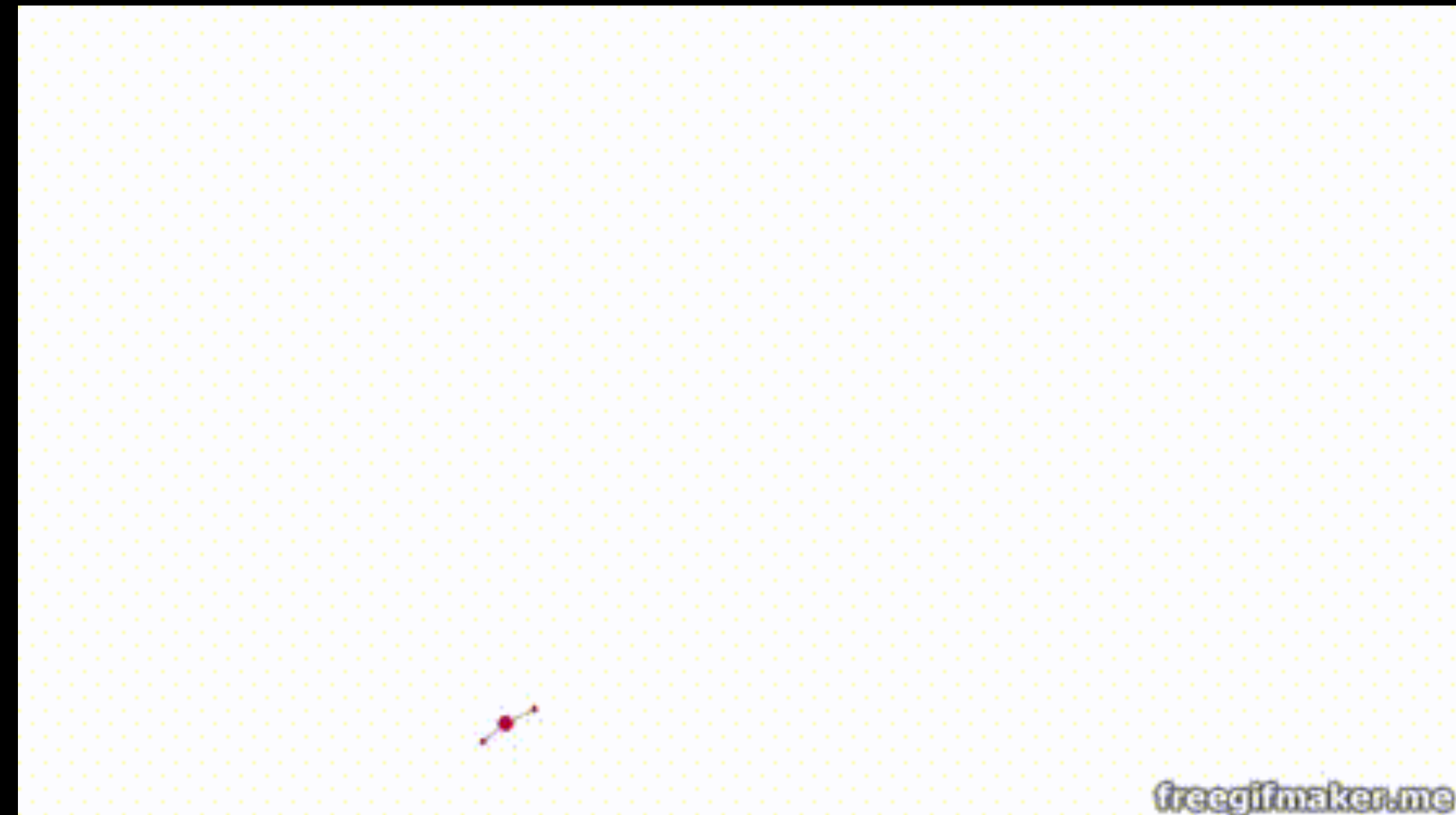
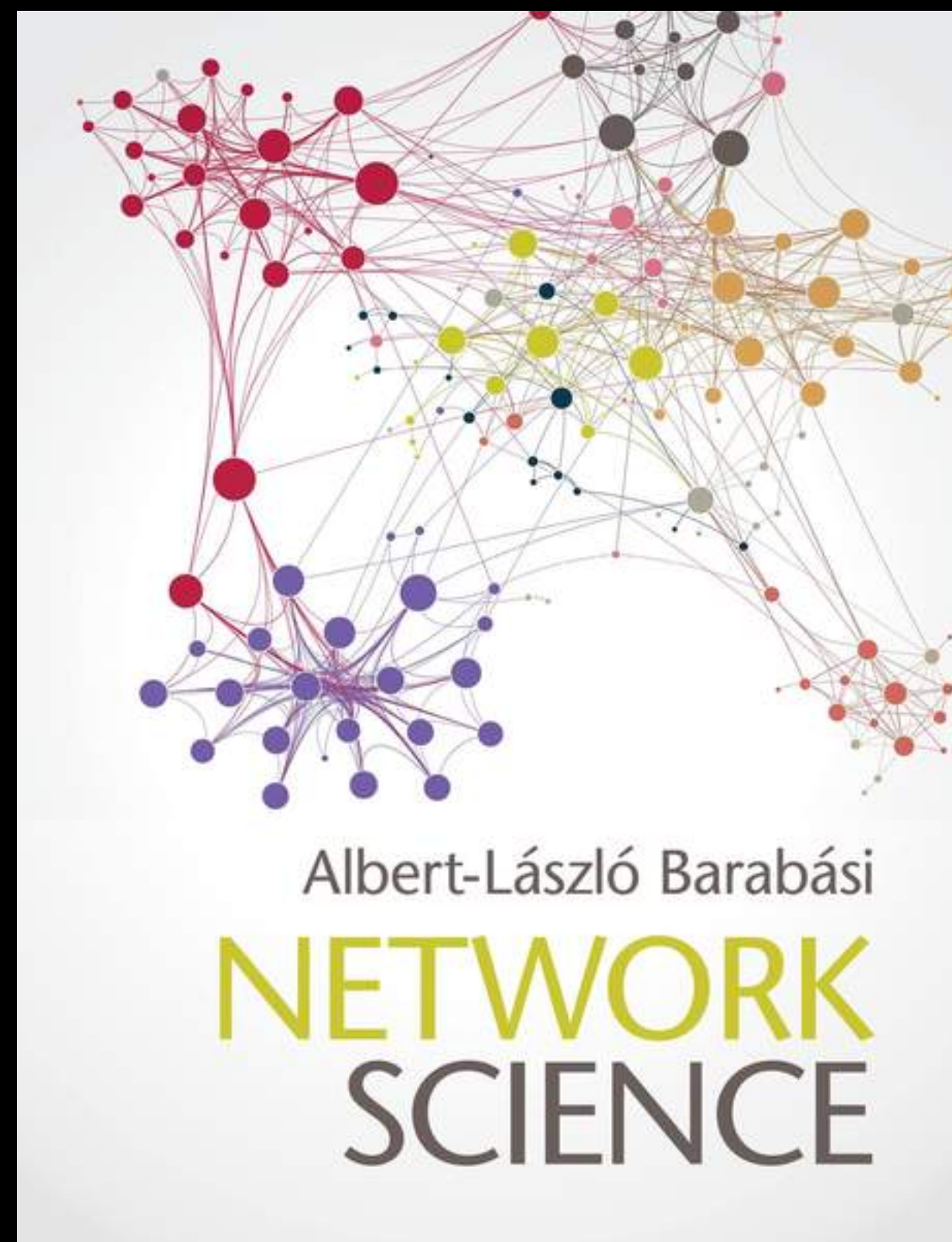
I will argue, in this paper, that the analysis of processes in interpersonal networks provides the most fruitful micro-macro bridge. In one way or another, it is through these networks that small-scale interaction becomes translated into large-scale patterns, and that these, in turn, feed back into small groups.

Sociometry, the precursor of network analysis, has always been curiously peripheral—invisible, really—in sociological theory. This is partly because it has usually been studied and applied only as a branch of social psychology; it is also because of the inherent complexities of precise network analysis. We have had neither the theory nor the measurement and sampling techniques to move sociometry from the usual small-group level to that of larger structures. While a number of stimulating and suggestive

¹ This paper originated in discussions with Harrison White, to whom I am indebted for many suggestions and ideas. Earlier drafts were read by Ivan Chase, James Davis, William Michelson, Nancy Lee, Peter Rossi, Charles Tilly, and an anonymous referee; their criticisms resulted in significant improvements.



Networks are flexible and exhibit emergent phenomena and equilibria



Preferential attachment

We can generate structurally realistic social networks

LLMs generate structurally realistic social networks but overestimate political homophily

Serina Chang^{1,*}, Alicja Chaszczewicz^{1,*}, Emma Wang¹,
Maya Josifovska^{1,2}, Emma Pierson³, Jure Leskovec¹

¹Department of Computer Science, Stanford University

²Department of Computer Science, University of California, Los Angeles

³Department of Computer Science, Cornell University

*These authors contributed equally.

Abstract

Generating social networks is essential for many applications, such as epidemic modeling and social simulations. Prior approaches either involve deep learning models, which require many observed networks for training, or stylized models, which are limited in their realism and flexibility. In contrast, LLMs offer the potential for zero-shot and flexible network generation. However, two key questions are: (1) are LLM's generated networks realistic, and (2) what are risks of bias, given the importance of demographics in forming social ties? To answer these questions, we develop three prompting methods for network generation and compare the generated networks to real social networks. We find that more realistic networks are generated with "local" methods, where the LLM constructs relations for one persona at a time, compared to "global" methods that construct the entire network at once. We also find that the generated networks match real networks on many characteristics, including density, clustering, community structure, and degree. However, we find that LLMs emphasize political homophily over all other types of homophily and *overestimate* political homophily relative to real-world measures.

1 Introduction

The ability to generate realistic social networks is crucial for many applications, when the true social network cannot be observed (e.g., for privacy reasons) or a realistic network is desired between hypothetical individuals. For example, in epidemic modeling, synthetic social networks are frequently used so that researchers can model the spread of disease based on who has come into contact with whom (Barrett et al., 2009; Block et al., 2020). Synthetic networks are also useful for simulating and analyzing social media platforms (Pérez-Rosés and Sebé, 2015; Sagduyu et al., 2018) and social phenomena, such as polarization and opinion dynamics (Dandekar et al., 2013; Das et al., 2014).

Deep learning approaches to network generation typically require training on many domain-specific networks (You et al., 2018), making it difficult to generalize to new settings where networks are not yet observed. Classical models for network generation require far less training, but these stylized models make rigid and unrealistic assumptions about how networks form. For example, Erdős-Rényi models assume that each edge forms with a uniform probability p (Erdős and Rényi, 1959). More realistic models, like small-world models (Watts and Strogatz, 1998) or stochastic block models (Holland et al., 1983), are still limited by a predefined, small set of numerical hyperparameters, missing the full complexity of real social interactions.

In contrast, generating social networks with large language models (LLMs) has the potential to address these limitations. LLMs possess zero-shot capabilities, enabling network generation without training. LLMs can also generate networks in a flexible manner, based on natural language descriptions of each person in the network. A key question, however, is whether LLMs can generate *realistic* social networks. On one hand, LLMs have demonstrated capabilities to realistically simulate human responses and interactions (Aher et al., 2023; Park et al., 2023; Argyle et al., 2023), suggesting that they may be able to generate realistic social networks as well. On the other hand, LLMs sometimes struggle with reasoning over graphs (Wang et al., 2023; Fatemi et al., 2024) and it is unclear if their language abilities generalize to structured objects like networks, so that they can reproduce structural characteristics of social networks such as low density and long-tailed degree distributions.

Furthermore, a central concern with using LLMs in social settings is bias. Prior works have shown that LLMs produce stereotyped descriptions of individuals based on their demographics (Cheng et al., 2023a,b) and skew towards the liberal opinions (Santurkar et al., 2023). These demographics,

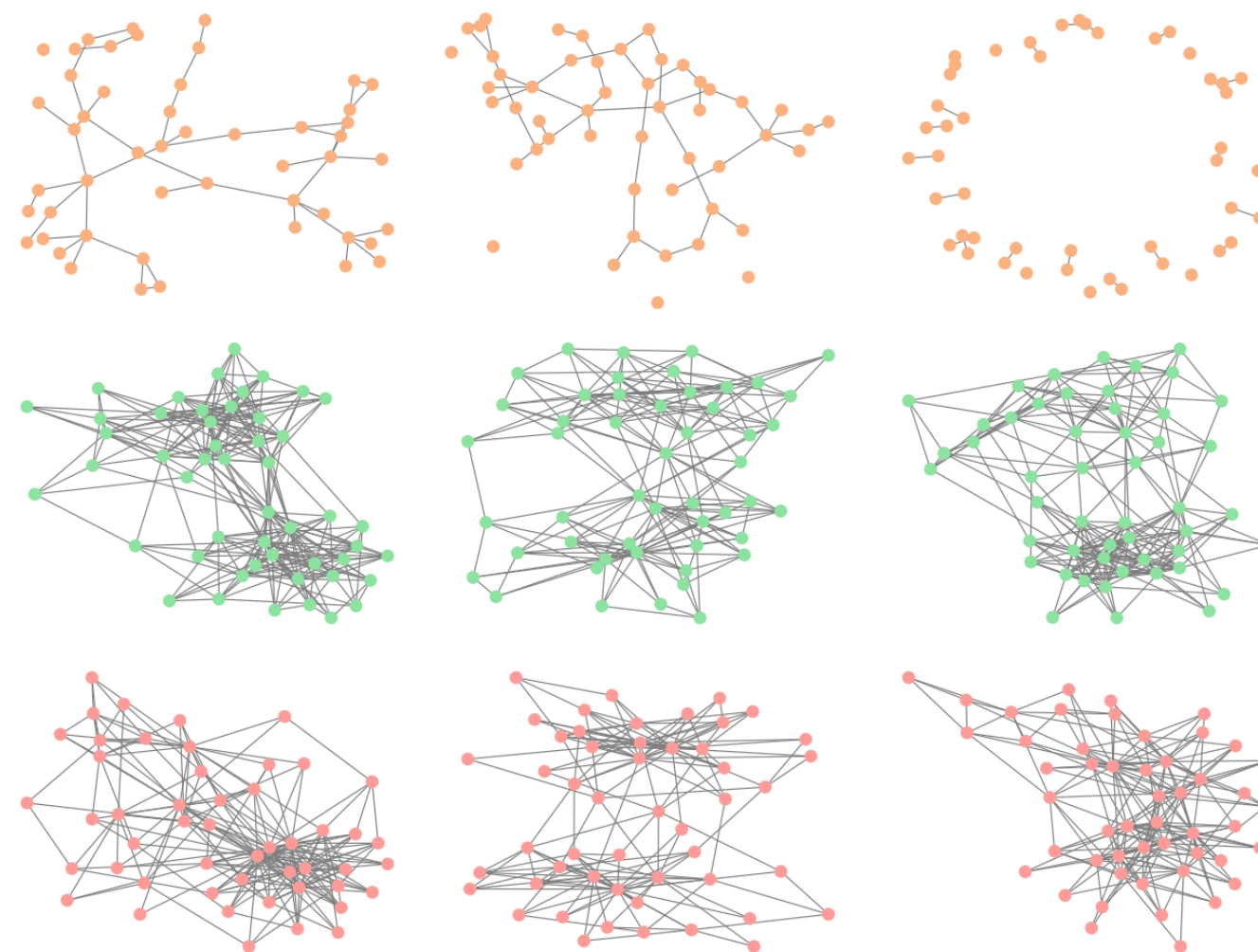


Figure 2: Generated social networks from different prompting methods: Global (top), Local (middle), Sequential (bottom).

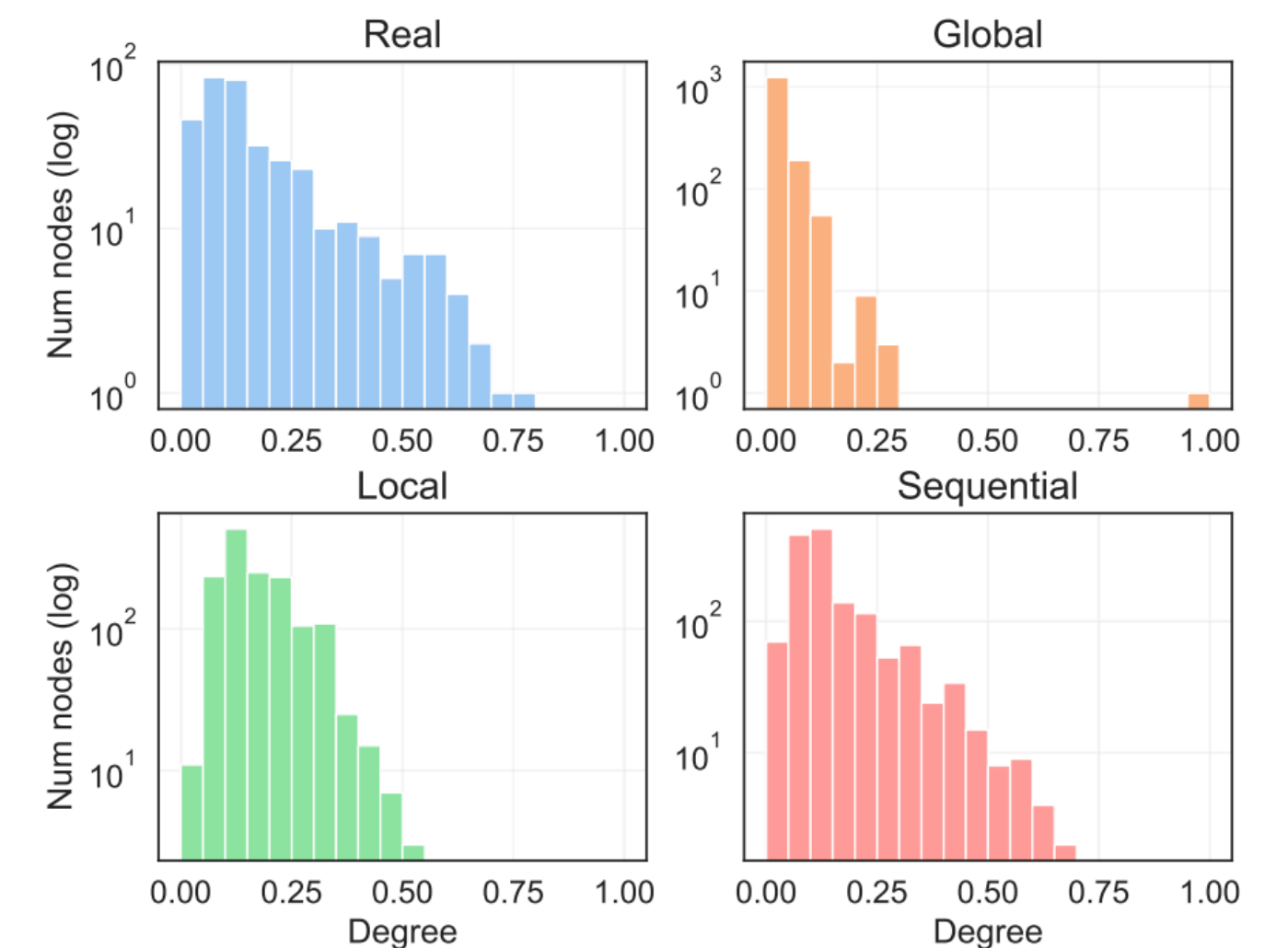


Figure 4: Degree distributions over real and generated social networks. For each set of networks, we pool degrees over nodes in the networks (Section 4).

Here, "realistic" could mean that we observe similar emergent phenomena

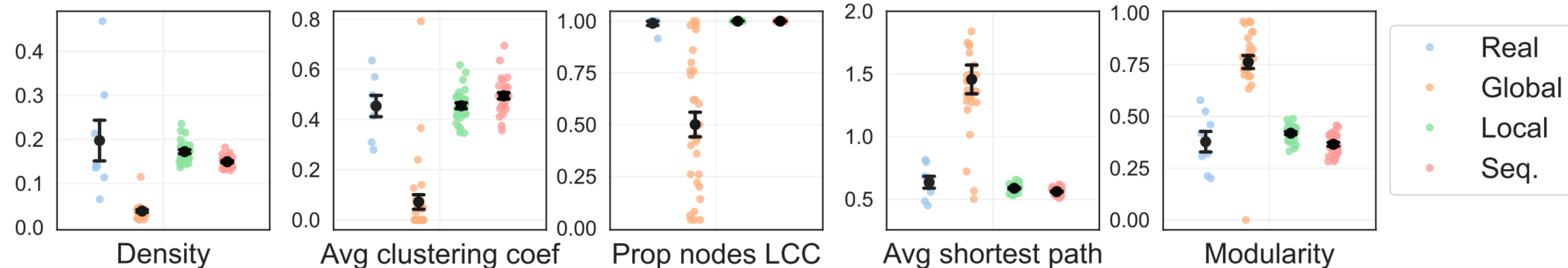


Figure 3: Graph-level metrics over real and generated social networks. We visualize mean and standard error (in black) and individual data points corresponding to each network.

Another angle: What if we generate the world in the same way we generate agent behaviors?

arXiv:2402.15391v1 [cs.LG] 23 Feb 2024

Genie: Generative Interactive Environments

Jake Bruce^{1,2}, Michael Dennis^{1,2}, Ashley Edwards^{1,2}, Jack Parker-Holder^{1,2}, Yuge (Jimmy) Shi^{1,2}, Edward Hughes¹, Matthew Lai¹, Aditi Mavalankar¹, Richie Steigerwald¹, Chris Apps¹, Yusuf Aytar¹, Sarah Bechtel¹, Feryal Behbahani¹, Stephanie Chan¹, Nicolas Heess¹, Lucy Gonzalez¹, Simon Osindero¹, Sherjil Ozair¹, Scott Reed¹, Jingwei Zhang¹, Konrad Zolna¹, Jeff Clune^{1,2}, Nando de Freitas¹, Satinder Singh¹ and Tim Rocktäschel^{1,2}

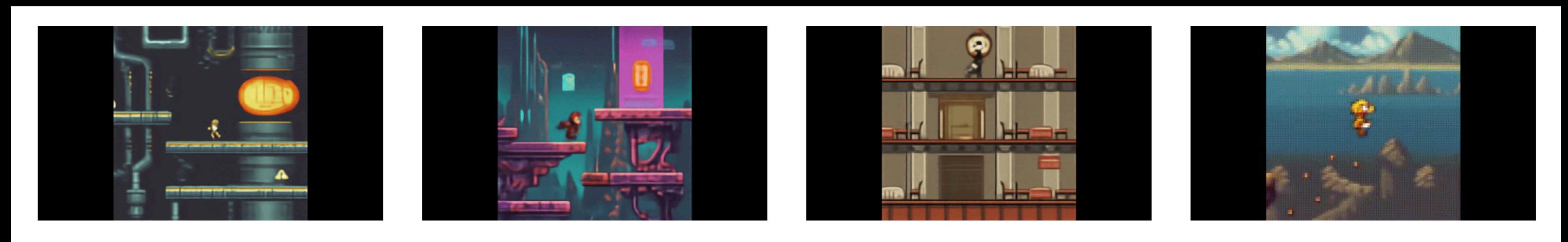
¹Equal contributions, ¹Google DeepMind, ²University of British Columbia



Figure 1 | **A whole new world:** Genie is capable of converting a variety of different prompts into interactive, playable environments that can be easily created, stepped into, and explored. This is made possible via a latent action interface, learned fully unsupervised from Internet videos. On the right we see a few generated steps for taking two latent actions. See more examples on our [website](#).

We introduce Genie, the first *generative interactive environment* trained in an unsupervised manner from unlabelled Internet videos. The model can be prompted to generate an endless variety of action-controllable virtual worlds described through text, synthetic images, photographs, and even sketches. At 11B parameters, Genie can be considered a *foundation world model*. It is comprised of a spatiotemporal video tokenizer, an autoregressive dynamics model, and a simple and scalable latent action model. Genie enables users to act in the generated environments on a frame-by-frame basis *despite training without any ground-truth action labels* or other domain-specific requirements typically found in the world model literature. Further the resulting learned latent action space facilitates training agents to imitate behaviors from unseen videos, opening the path for training generalist agents of the future.

Keywords: Generative AI, Foundation Models, World Models, Video Models, Open-Endedness



References

- M. J. Salganik, P. S. Dodds, D. J. Watts, Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science* 311, 854-856 (2006).
- Park, J.S., Popowski, L., Cai, C.J., Morris, M.R., Liang, P., & Bernstein, M.S. (2022). Social Simulacra: Creating Populated Prototypes for Social Computing Systems. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22). Association for Computing Machinery, New York, NY, USA.
- R. Thaler, Mental Accounting and Consumer Choice. *Marketing Science* 4, 199-214 (1985).
- T. C. Schelling, Dynamic models of segregation. *J. Math. Sociol.* 1, 143-186 (1971).
- J. von Neumann, O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, 1944).
- L. P. Argyle et al., Out of one, many: Using language models to simulate human samples. *Political Analysis* 31, 337-355 (2023).
- A. Ashokkumar, L. Hewitt, I. Ghezze, R. Willer, "Predicting Results of Social Science Experiments Using Large Language Models" (2024).

References

- J. J. Horton, "Large language models as simulated economic agents: What can we learn from homo silicus?" (2023).
- R. Louie, A. Nandi, W. Fang, C. Chang, E. Brunskill, D. Yang, Roleplay-doh: Enabling Domain-Experts to Create LLM-simulated Patients via Eliciting and Adhering to Principles. Preprint (2024).
- Park, J.S., Popowski, L., Cai, C.J., Morris, M.R., Liang, P., & Bernstein, M.S. (2022). Social Simulacra: Creating Populated Prototypes for Social Computing Systems. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22). Association for Computing Machinery, New York, NY, USA.
- O. Shaikh, V. Chai, M. J. Gelfand, D. Yang, M. S. Bernstein, Rehearsal: Simulating Conflict to Teach Conflict Resolution, in Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24), Honolulu, HI, USA, May 11-16, 2024.
- J. S. Park, J. C. O'Brien, C. J. Cai, M. R. Morris, P. Liang, M. S. Bernstein, Generative agents: Interactive simulacra of human behavior, in Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (ACM, 2023).

References

- C. Qian, W. Liu, H. Liu, N. Chen, Y. Dang, J. Li, C. Yang, W. Chen, Y. Su, X. Cong, J. Xu, D. Li, Z. Liu, M. Sun, ChatDev: Communicative Agents for Software Development, in Proceedings of the 2024 Annual Conference of the Association for Computational Linguistics (ACL 2024).
- P. Jansen, M.-A. Côté, T. Khot, E. Bransom, B. Dalvi Mishra, B. P. Majumder, O. Tafjord, P. Clark, DISCOVERYWORLD: A Virtual Environment for Developing and Evaluating Automated Scientific Discovery Agents. Preprint (2024).
- J. Li, S. Wang, M. Zhang, W. Li, Y. Lai, X. Kang, W. Ma, Y. Liu, Agent Hospital: A Simulacrum of Hospital with Evolvable Medical Agents. Preprint (2024).
- A. Rosinol, A. Violette, M. Abate, N. Hughes, Y. Chang, J. Shi, A. Gupta, L. Carlone, Kimera: from SLAM to Spatial Perception with 3D Dynamic Scene Graphs. *Int. J. Robot. Res.* 40, 1510-1546 (2021).
- M. S. Granovetter, The Strength of Weak Ties. *Am. J. Sociol.* 78, 1360-1380 (1973).
- A. L. Barabási, *Network Science* (Cambridge Univ. Press, Cambridge, 2016).
- A. L. Barabási, R. Albert, Emergence of Scaling in Random Networks. *Science* 286, 509-512 (1999).

References

- S. Chang, A. Chaszczewicz, E. Wang, M. Josifovska, E. Pierson, J. Leskovec, LLMs generate structurally realistic social networks but overestimate political homophily. Preprint (2024).
- J. Bruce et al., Genie: Generative Interactive Environments. Preprint (2024).



The image shows a top-down view of a simulated environment, likely a campus or park. It features several interconnected buildings with various rooms, including offices, classrooms, a library, and a dining area. Each room contains furniture like desks, chairs, and bookshelves. Numerous small, stylized human figures (agents) are scattered throughout the environment, each with a speech bubble containing a two-letter code (e.g., LW, RP, AC, AB, IR, GR, CG, FL, HJ, WS, JL, KM, AS, YY, JM, TT, CO, TM, ML, EL). The environment is surrounded by green grass, trees, and a central dirt path. The overall style is a colorful, pixelated aesthetic.

CS 222: AI Agents and Simulations

Stanford University

Joon Sung Park