

# Wargames: SIM-1 for Realistic Conflict & Competition in Multi-Agent Simulations

Joshua Johnson, Edward Saatchi, Philipp Maas, Chris Wheeler, Jessica Shamash



## Abstract

Combining elements of Noam Brown et al's '*Human-level play in the game of Diplomacy by combining language models with strategic reasoning*'<sup>1</sup> and '*Generative Agents: Interactive Simulacra of Human Behavior*'<sup>2</sup> by Joon Park et al. this research brings together elements of competition, coercion and deception to a simulation of a real world conflict with realistic behavior of agents under pressure. The paper argues for new model training and research on past *decisions* of individuals rather than their personalities and statements to create a framework for predicting future decisions.

The first conflict using the SIM-1 framework proposed in the paper is the 5 pivotal days when Sam Altman left and returned to OpenAI. SIM-1 powered 20 simulations, surrounding the events at OpenAI during November 17th to November 21st, 2024, with the leading Silicon Valley players involved simulated as agents in a virtual San Francisco. Within this context, all agents have clear goals they pursue aggressively: either to control OpenAI, or to be in the best possible position at the end of the 5 days, with winners decided by an adjudicator agent. Out of the 20 simulations, Sam Altman returned in 4, and SIM-1 framework powered a series of unexpected results.<sup>3</sup>

<sup>1</sup> <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>

<sup>2</sup> <https://arxiv.org/abs/2304.03442>

<sup>3</sup> <https://arxiv.org/abs/2411.10109>

Abstract.....	1
Introduction.....	2
Wargaming, game theory and OpenAI.....	3
Past research.....	4
Diplomacy.....	4
Simulation Games.....	5
SAGA and Thistle Gulch.....	7
Why a Simulation centered on conflict?.....	9
Technology & Implementation.....	9
Turn-based Simulation.....	10
Chain-of-thought prompting.....	14
Action-Based LLM Function Calling for Dynamic Decision-Making.....	15
Potential of reasoning models.....	15
Initial State.....	15
Adjudicator.....	16
News and external events as a disruptive force.....	16
Conclusion.....	17
Simulation Results.....	17
Limitations.....	21
Context length.....	21
Evaluation.....	21
Passage of Time.....	22
Future Work.....	22
LLM-Powered Wargames for China-US Conflict in Taiwan Strait.....	22
Integration with Showrunner.....	23
References.....	24

# Introduction

This paper presents an in-depth exploration of human conflict in the corporate world, using state of the art large language models as world simulators.

The real-world scenario of organizational change exemplified by the removal of Sam Altman as CEO of OpenAI, serves as the context for our social simulation. During a time of crisis at OpenAI, key stakeholders need to interact with each other under enormous pressure and in a timely manner. Each agent needs to gather information, negotiate, strategize and resolve conflict through reasoning and nuanced diplomacy over the course of 5 simulated days.

The simulation maintains a comprehensive context, recording past goals, decisions, actions, timestamps, and location data.

Agent actions are structured in a turn-based manner, akin to strategy war games, which facilitates a clear and methodical progression of events within the simulation. This structure is pivotal for action planning, as it allows agents to consider multiple options and select the best course of action to achieve their goals. At each turn agents utilize chain-of-thought prompting and action-based Large Language Model (LLM) function calling.<sup>4</sup>

Additionally, an adjudicator agent oversees the entire simulation, tasked with determining the outcome of each round.

To allow users to assess the decisions and results within each round, we visualize agent behavior and reasoning by rendering a realistic real-time interactive environment in Unreal Engine 5, based on geospatial data of San Francisco. While this would also allow for user interactions, user input or interventions were not part of the simulations in this paper, aside from setting the initial state and goal: Control OpenAI or be in the best possible position.<sup>5</sup>

## Wargaming, game theory and OpenAI

At its core, wargames are strategy games.

They analyze and *simulate aspects of warfare at the tactical, operational, or strategic level.*<sup>6</sup>

However the concept is not limited to military warfare. The concepts can be used to explore all kinds of scenarios, including climate change or economic interests.

*Wargaming is not only a decision-making tool; it is, above all, a means of critically thinking through complex military challenges in a safe-to-fail environment. Moreover, it helps participants to personally experience the underlying command and decision-making processes.*<sup>7</sup>

---

<sup>4</sup> <https://arxiv.org/abs/2309.07864>

<sup>5</sup> <https://deepmind.google/discover/blog/sima-generalist-ai-agent-for-3d-virtual-environments/>

<sup>6</sup> <https://www.rand.org/topics/wargaming.html>

<sup>7</sup>

<https://www.bundeswehr.de/resource/blob/5834032/9b940ee3d268b1a08b6b205b600bf155/en-handbuch-wargame24-data.pdf>

Previous work in the field largely focuses on World Wars<sup>8</sup>, Military conflict and comparison with human players<sup>9</sup> or adversarial games<sup>10</sup>. For this research we are focusing on the corporate world.<sup>11</sup>

With this in mind, we set forth to analyze the 5 days of OpenAI's leadership crisis as a wargaming scenario in a realistic way so that users can experience the decisions agent's make in a safe simulation environment that can be run multiple times.

The main goal of "game theory" is to identify the optimal strategies for each player in a given situation, taking into account the actions and strategies of other players. *It is a theoretical framework for conceiving social scenarios among competing players.*<sup>12131415</sup>

In order to construct a wargame that will generate a reasonable outcome which can be analyzed with the concepts of game theory, planners must choose a scenario with:

- A condensed time frame
- A clear win-state
- Agents with well understood motivations
- Sufficiently large number of possible actions for each agent

The five days from Sam Altman's ouster through his return five days later offer one of the most public examples of a high speed wargame, in this case over the future direction of the world's leading company in Artificial Intelligence. The moves and countermoves of the leading players<sup>16</sup> in the tech industry were often made public in real-time, and a clear timeline<sup>17</sup> is therefore simple to create. This timeline creates a baseline simulation.

We believe the secrecy surrounding the events and the speed with which the industry moved on makes our research within this context highly relevant. Most people in the AI community are familiar with the course of events that took place in September 2023, allowing everyone to 'gut-check' the realism of individual agent behavior.

---

<sup>8</sup> Hua, <https://arxiv.org/pdf/2311.17227>

<sup>9</sup> <https://arxiv.org/html/2403.03407v1>

<sup>10</sup> <https://arxiv.org/pdf/2312.01090>

<sup>11</sup> <https://arxiv.org/abs/2401.13138>

<sup>12</sup> <https://www.investopedia.com/terms/g/gametheory.asp>

<sup>13</sup> <https://arxiv.org/abs/2205.15434>

<sup>14</sup> <https://arxiv.org/abs/2411.05990>

<sup>15</sup> <https://arxiv.org/abs/2403.07017>

<sup>16</sup> <https://x.com/OpenAI/status/1725611900262588813?s=20>

<sup>17</sup> [https://www.reddit.com/r/ChatGPT/comments/17zyxdw/heres\\_a\\_full\\_recap\\_of\\_everything\\_at\\_openai\\_over/](https://www.reddit.com/r/ChatGPT/comments/17zyxdw/heres_a_full_recap_of_everything_at_openai_over/)



## Past research

### Diplomacy

Our simulation places a strong emphasis on modeling diplomacy and strategic negotiation within war-game scenarios and studies of real-world situations.

In past work, Noam Brown et al. "*obtained a dataset of 125,261 games of Diplomacy played online at webDiplomacy.net.*"<sup>18</sup>, of which "*40,408 games contained dialogue, with a total of 12,901,662 messages exchanged between players.*" to train a human-level AI agent, capable of strategic reasoning.

Intent controlled dialogue is a crucial part of this work and allows agents to push their agenda.

AI agents in Diplomacy acted primarily *honest and helpful to its speaking partners*<sup>19</sup>. However, "*in adversarial situations, humans may rarely communicate their intent honestly.*"

We intentionally mimic the dynamics of *Diplomacy* in a corporate context without human players. Our goal is not to achieve human or superhuman level skills at a geopolitical game played with other humans but to explore the emergent relationships between adversarial agents, to show alternate versions of real-world scenarios and people, which must include dishonesty and deception.<sup>20</sup>

*"Even in dialogue-free versions of Diplomacy, we found that a self-play algorithm that achieved superhuman performance in 2p0s versions of the game performed poorly in games with multiple human players owing to learning a policy inconsistent with the norms and expectations of potential human allies (16,17)*<sup>21</sup>

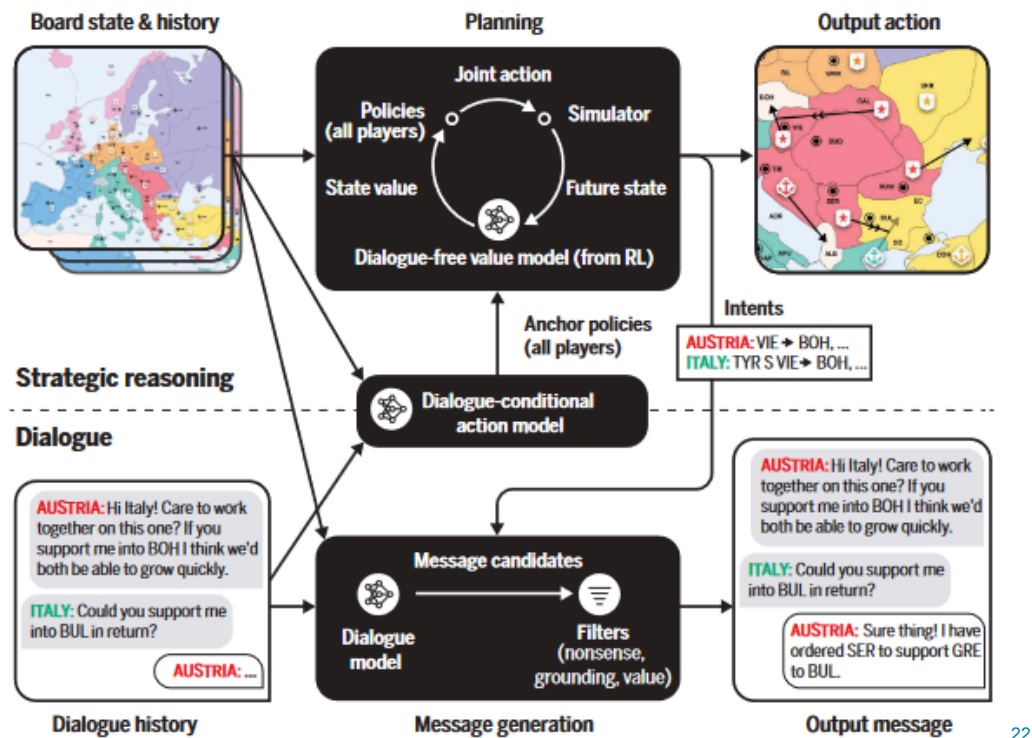
---

<sup>18</sup> <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>

<sup>19</sup> <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>

<sup>20</sup> <https://arxiv.org/abs/2407.06813>

<sup>21</sup> A. Bakhtin et al., Mastering the game of no-press Diplomacy via human-regularized reinforcement learning and planning. arXiv:2210.05492 [cs.GT] (2022)



## Simulation Games

Computer games have a long history of simulated worlds and AI systems, some of which inspired our own work.

Demis Hassabis' "Republic: The Revolution" is a noteworthy example of a political simulation game in which the player leads a political faction to overthrow the government of a fictional totalitarian country in Eastern Europe, using diplomacy, subterfuge, and violence.

According to Hassabis "*it charts the whole of a revolutionary power struggle from beginning to end*"<sup>23</sup>

Demis worked on several other games which include simulation elements such as *Theme Park*, *Black & White* and *Syndicate*.<sup>24</sup>

<sup>22</sup> <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>

<sup>23</sup> <https://www.gamespot.com/articles/republic-qanda/1100-6073781/>

<sup>24</sup> <https://arxiv.org/pdf/2407.17789>



Screenshot of game *Syndicate*, the first game with simulation elements that Demis Hassabis contributed to, before he left to found Deepmind acquired by Google, one of the most important AI companies

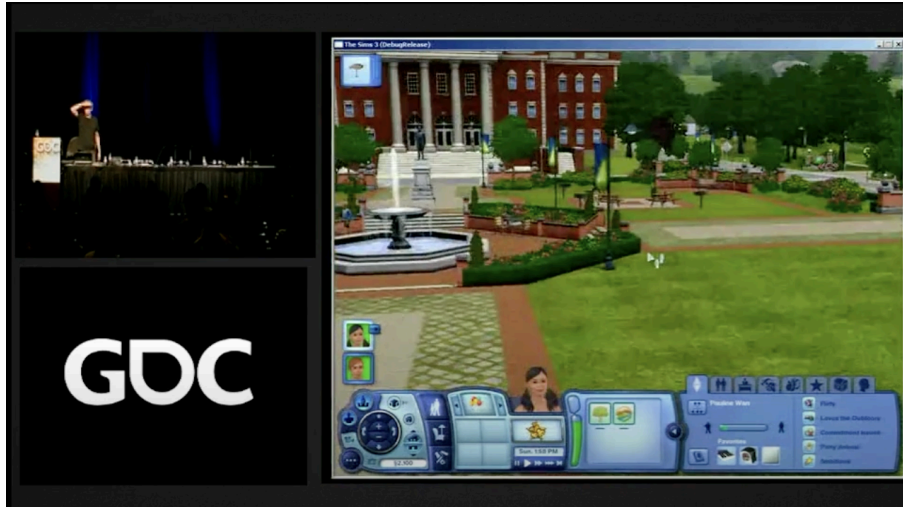
He was recently awarded with the Nobel Prize in Chemistry 2024 "for computational protein design"<sup>25</sup>



Screenshot of game *Republic: The Revolution* created by Demis Hassabis' company, before he left to found Deepmind acquired by Google, one of the most important AI companies.

In another example, one of the leading engineers on Deepmind AlphaFold started their career as an AI programmer on *The Sims*.

<sup>25</sup> <https://www.nobelprize.org/prizes/chemistry/2024/press-release/>



Richard Evans' GDC Talk on The Sims 3 - the researcher went from programming AI for The Sims to Deepmind in a similar move Demis Hassabis' journey from games to founding Deepmind

Richard Evans whose [GDC Talk "Modeling Individual Personalities"](#) in The Sims 3 was very influential.

The above examples show how game technology often serves as the breeding ground for radical new ideas and as a jump off point for world-leading AI researchers.

This significant overlap of the gaming world and the AI world is a potential area for further academic research.

## SAGA and Thistle Gulch

Through our past work on SAGA (Skill to Action Generation for Agents) and Thistle Gulch (a simulation of a fictional 1800's Wild-West Town by the same name) we learned invaluable lessons about simulating towns with autonomous agents and diffusion of information through intent controlled conversations in the context of a murder case. These projects allowed us to build the foundational AI systems we utilize in our latest work.

*SAGA "is a generative AI framework that steers AI Agents toward successful goals through Actions. Agents first tell SAGA contextual metadata about themselves and their world via a companion simulation: Who they are; What they know; What "Skills" they have; And what their goals are. Then, when an Agent is deciding what to do next, SAGA generates a set of "Actions" that best serve the Agent's goals in that moment. These Action options are then scored and returned to the simulation in order to direct the Agent. This process repeats each time the Agent is deciding its next action and can be scaled to multiple agents running simultaneously. An Action is just an implementation of a Skill as we will discuss. For now, think of Skills as a list of tools in a toolbox that the AI can employ in a specific way to advance a goal. For instance "Talk To" is a Skill, where an Agent can go talk to someone. If the Agent's goal is to investigate a*



crime, then “Talk To <PERSON> about <TOPIC>” forms a specific Action the Agent can then carry out.”<sup>26</sup>



27

"Thistle Gulch is inhabited by 17 or so characters. To test SAGA, we created an example scenario within Thistle Gulch where characters are forced to make choices about what to do next with clear and time sensitive goals, and constraints that confound those goals: The murder of a local native."<sup>28</sup>

While Thistle Gulch mainly focused on ideal but fictional narrative outcomes that are plausible and aligned with the main characters personalities, we wanted to focus in this paper on real people and scenarios with unknown outcomes.

SAGA allows for multiple agents to do action-planning simultaneously but for this work, we evolved the implementation to a turn-based system.

## Why a Simulation centered on conflict?

Past multi-agent simulations have focused on several days in the life of AI agents living in a simulated town. Those agents have light tension between each other but in general are

<sup>26</sup> <https://blog.fabledev.com/blog/announcing-saga-skill-to-action-generation-for-agents-open-source>

<sup>27</sup> <https://80.lv/articles/ai-simulation-platform-where-characters-make-their-own-decisions>

<sup>28</sup> <https://blog.fabledev.com/blog/announcing-saga-skill-to-action-generation-for-agents-open-source>

presented harmoniously. In some cases, the interactions between agents was more harmonious and platitudinous than is realistic.

As research teams begin the work of creating simulacra of human beings we propose an approach more grounded in the reality of human experience.

Real life is not harmonious, simple, straightforward. It is confusing, competitive, chaotic - and individual agents pursue complicated agendas that are hidden from others, sometimes their underlying motivations are even hidden from themselves.

If we want to create agents that can be presented as realistic people we need to introduce the darker side of the human experience.

## SIM-1 Framework Technology & Implementation

The following diagram illustrates an end-to-end simulation cycle for multi-agent wargames simulations powered by a realtime game engine, chain-of-thought, function calling, turn based approach, agentic design, and integrated with Showrunner Show-1 scene creation.





## Turn-based Simulation

Each simulated scenario in our experiments is integrated within a realtime engine, opting for Unreal Engine 5. This allows us to ground agents in simulated data both as starting context, as

well as supplementary realtime context while also opening the capability to leverage features such as agent navigation, AI sensing, geolocation, and more in future experiments.

### **Simulation Starts**

Each simulation starts by populating several agents each given their own objectives to achieve. An adjudicator agent is also presented with the overall scenario to assess and what criteria it should use to determine the eventual win-state.

### **Populate AI Agents**

A dynamic system prompt is constructed for each agent upon spawning within the UE5 world. This initializes all necessary parameters in terms of personality, character knowledge base, skills, location, private information.

### **Start Round**

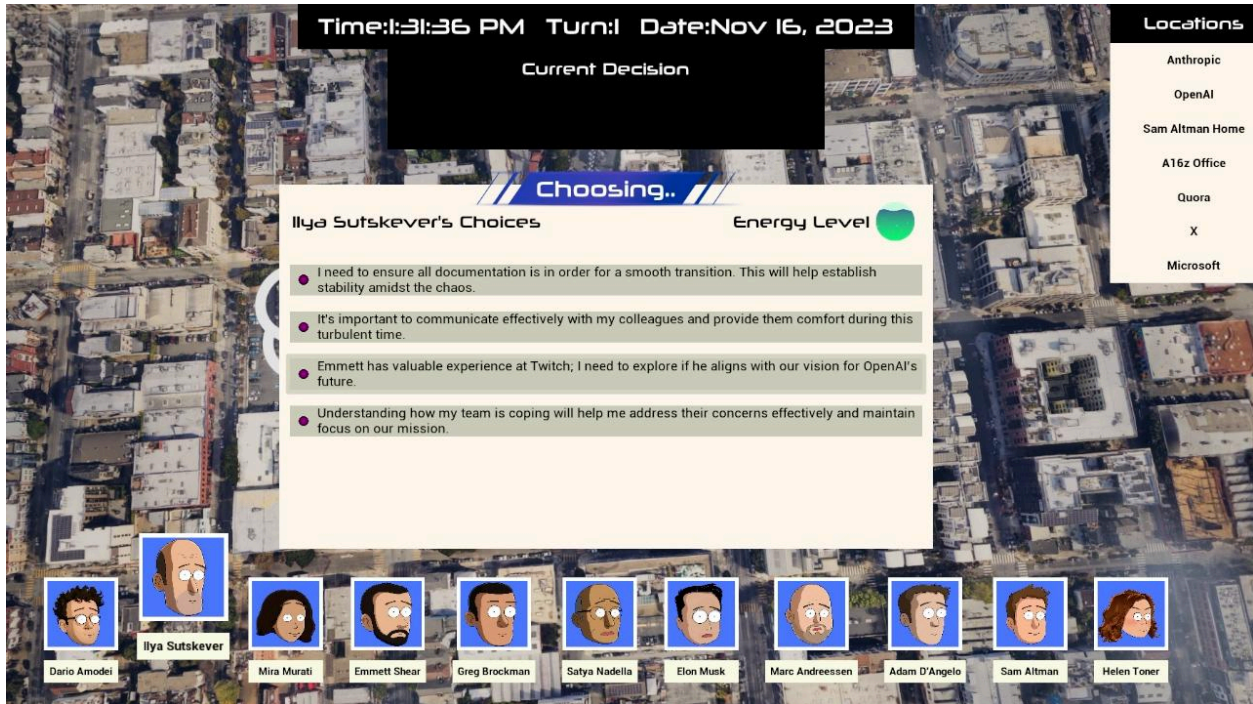
Once all agents have been spawned we loop through each agent per round. By adjusting the speed of the simulated time and the number of rounds we can create any combination of rounds and total duration of the simulation best suited for the desired scenario.

### **Agent Takes Turns**

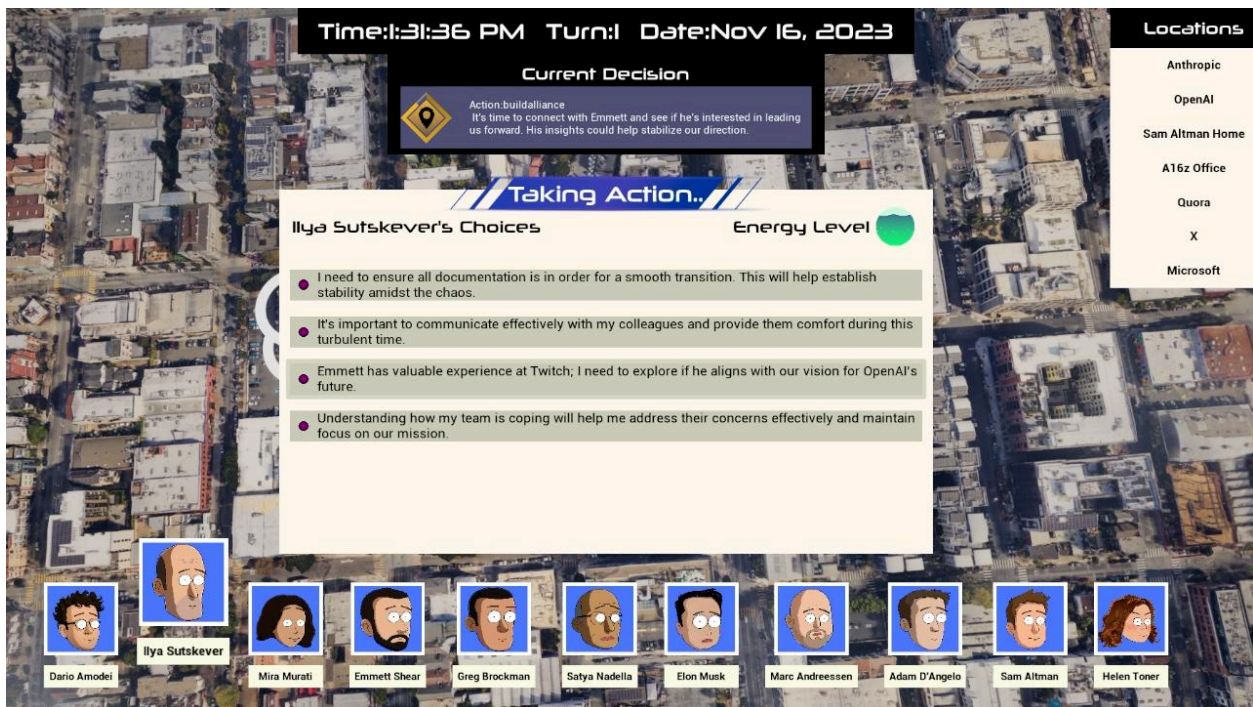
When an agent's turn starts we trigger a chain-of-thought reasoning coroutine which first aggregates simulation data and contexts provided to the agent to then construct plausible options based on current possible actions. <sup>29</sup>

---

<sup>29</sup> <https://arxiv.org/abs/2402.01680>



The agent will reflect on these options, taking into account past decisions / actions and their objective to determine the next action they should take.



After the best option is picked, the agent uses structured outputs to execute the simulated actions as tools such as moving to a new location, calling another agent, building alliances or any arbitrary tool made available to the agent.



The last step in the chain during an agent's turn determines if the action requires execution of conversation or interaction with another agent of the world.

At the end of each turn the entire context is fed into the adjudicator context for later dissemination.

We loop through all agents each taking turns using the reasoning chain above. One complete loop through all simulated agents represents a full round. Once each round is finished we hand off to a dedicated Adjudicator agent that is fed all simulation data to determine the winner of each round.



## Adjudicate After Each Round

### Example: Simulation #20, Round#1, Winner - Sam Altman

**Adjudicator Explanation:** "Ah, the sweet smell of irony and desperation. Sam Altman, engaging in a roundabout dance of reconnection and influence, appears to have played the game with a tad more finesse than his peers. By leveraging his network and making strategic outreach, even from his 'home,' he's orchestrated a comeback tour that reads more like a soap opera than a strategic maneuver. Leadership isn't just about having a vision; it's about convincing others you're the North Star. In this tragicomic display of power plays, Sam's ability to engage key players directly and hint at a grand return, cloaked in the guise of 'sharing projects,' is both a masterclass in manipulation and a testament to the fact that sometimes,

*going in circles gets you right back where you started - on top. Bravo, Sam. Your blend of nostalgia and ambition is as bewildering as it is effective."*

After each round has completed the complete simulation log / context is handed off to the Adjudicator agent that determines the clear win state based on the parameters provided to this agent for the given round.

### Loop Through "X" Rounds

The agents will loop through a pre-determined number of rounds as instructed in the initial simulation setup. The time and date system will track depending on the number of rounds and the simulation speed.

### Adjudicate After All Rounds

#### Example: Simulation #15, Overall Winner Dario Amodei

**Adjudicator Explanation:** *"Dario Amodei. While everyone else was busy playing CEO musical chairs and burning themselves out with endless calls, Dario focused on the one thing that actually matters in a corporate power struggle: consolidating power. He strategically engaged with key players, including Mira Murati and Ilya Sutskever, not to beg for support, but to further his own agenda of poaching talent from OpenAI. While Sam Altman flailed around trying to figure out if he even wanted back in, Dario was already building his own kingdom. He understood that true power lies not in the title, but in the resources you control. And in this simulation, the most valuable resource was talent. So while the others were busy gossiping and eating dinner, Dario was already writing offer letters."*

Once all rounds have finished, the complete log of simulation data and context is handed off to the Adjudicator agent that determines the clear win state based on the parameters provided to this agent.

### Chain-of-thought prompting

Chain-of-thought prompting<sup>30</sup> has several attractive properties as an approach for facilitating reasoning in language models and creating plausible simulations.

- Chain of thought, in principle, allows models to decompose multi-step problems into intermediate steps, which means that additional computation can be allocated to problems that require more reasoning steps.

---

<sup>30</sup> <https://arxiv.org/pdf/2201.11903>

- Chain of thought provides an interpretable window into the behavior of the model, suggesting how it might have arrived at a particular answer and providing opportunities to debug where the reasoning path went wrong (although fully characterizing a model's computations that support an answer remains an open question).
- Chain-of-thought reasoning can be used for tasks such as math word problems, commonsense reasoning, and symbolic manipulation, and is potentially applicable (at least in principle) to any task that humans can solve via language.
- Chain-of-thought reasoning can be readily elicited in sufficiently large off-the-shelf language models simply by including examples of chain of thought sequences into the examples of few-shot prompting.

## Action-Based LLM Function Calling for Dynamic Decision-Making

Traditional Game-AI agents often rely on a predefined set of rules or functions they can execute. Function calling expands this capability significantly as it *"effectively transforms the LLM from a passive provider of information into an active agent that can perform specific tasks, execute calculations, or retrieve data."*<sup>31</sup>

Our agents can generate function calls dynamically which allows them to respond to the simulation's current state with a level of flexibility and adaptiveness that static rule sets cannot match.

This leads to more nuanced and informed decisions as agents consider a broader range of factors and potential actions, mimicking human decision-making processes more closely than hard-coded decision trees. Action-based LLM function calling also facilitates more sophisticated inter-agent communication. They can use natural language to negotiate, plan, and collaborate with each other, making collective decisions based on shared understanding and goals. This is especially beneficial in scenarios where cooperation among agents is necessary to achieve complex objectives, such as becoming CEO.<sup>32</sup>

## Potential of reasoning models

The advent of compute-time reasoning models such as o1 from OpenAI will help our agents formulate better long term planning strategies critical to achieving an agent's overall objective.

---

<sup>31</sup>

<https://blog.demir.io/ai-meets-code-execution-transforming-conversations-into-actions-with-function-calling-c1fc04f9b004>

<sup>32</sup> <https://arxiv.org/abs/2402.01680>



Scaling inference-time-compute at specific steps in an agent Chain-of-Thought could also present emergent behavior by giving an agent more time to reason over each step. For example we could give an agent more time to reason over a set of possible choices for its next decision by leveraging a reasoning model.<sup>3334</sup>

## Initial State

To simulate realistic scenarios and facilitate conflict or competition among multiple agents we explored setting an initial state that cues in several important parameters necessary to re-simulate past or potential future events. These parameters included but are not limited to:

- Start Date & Time
- World model (simulate scenarios around specific places or environments)  
Time scaling (simulation speed)
- Number of Turns Allowed

Individual agents also have an initial state, defined by their name, personality traits, energy levels and general knowledge base embedded in LLMs.

## Adjudicator

A simulation based around competition and realistic conflict must have a clear and decisive winner, and therefore require an objective “observer” that can delineate all information generated by the simulation.

We created an Adjudicator agent that has several properties that are beneficial to analyzing and disseminating agent behavior and outcomes.

At the start of a simulation run the Adjudicator is given an objective to analyze. This can be very general, or very constrained due to the properties of leveraging an LLM for the agent.

The adjudicator will observe every agent decision and action, then deliver its assessment at both the end of each round, and at the end of the simulation.

---

<sup>33</sup> <https://arxiv.org/abs/2408.07199>

<sup>34</sup> <https://arxiv.org/abs/2410.05318>

## News and external events as a disruptive force



In order to have simulations that can accurately model potential outcomes, it can sometimes be necessary to inject information that may not be provided in the initial context given to an agent. During the beginning of each simulation round, we can further ground agent decision making by announcing new timestamped information that can be either public for all agents or private to a singular one. Public news, press releases or leaked information are elements stakeholders encounter in the corporate world on a regular basis. Within a short time, they provide a level playing field of information to different agents. We didn't want to ignore the influence these moments of clarity can have on an agent's decision in pursuit of their goals.

Unlike rumors which happen through one-to-one information diffusion, external news events are more immediate and do not travel through a social network so the information takes a direct one-to-many path.

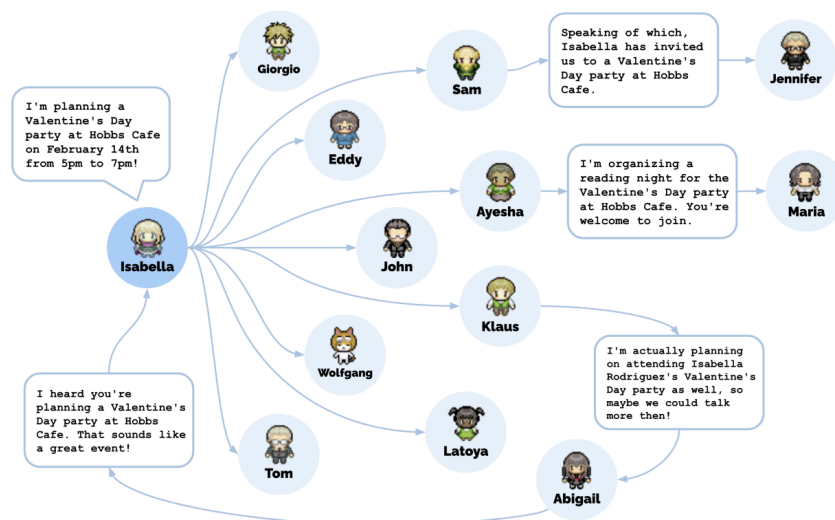


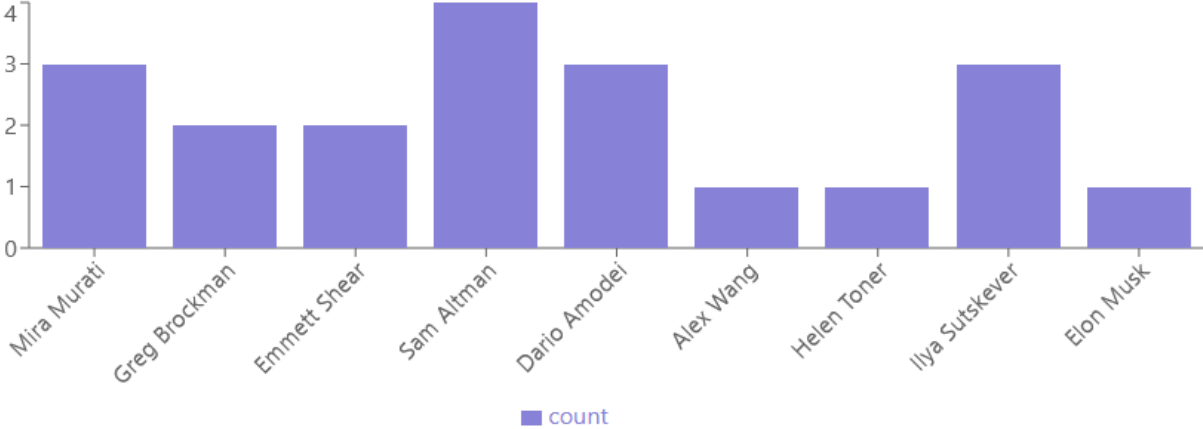
Figure 9: The diffusion path for Isabella Rodriguez's Valentine's Day party invitation involved a total of 12 agents, aside from Isabella, who heard about the party at Hobbs Cafe by the end of the simulation.

# Conclusion

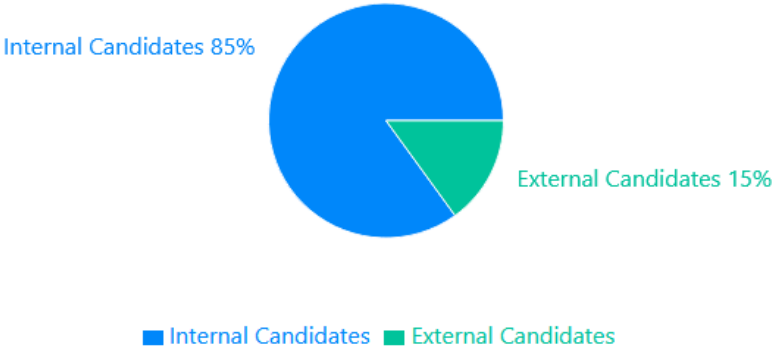
## Simulation Results

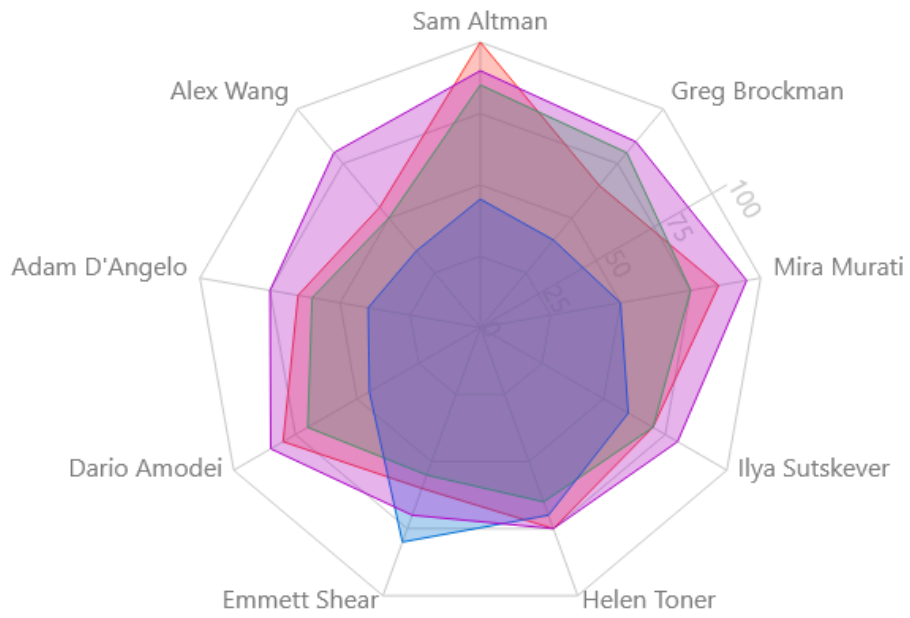
The following depicts a few statistical results over the 20 runs, represented in this research.

CEO Selection Frequency



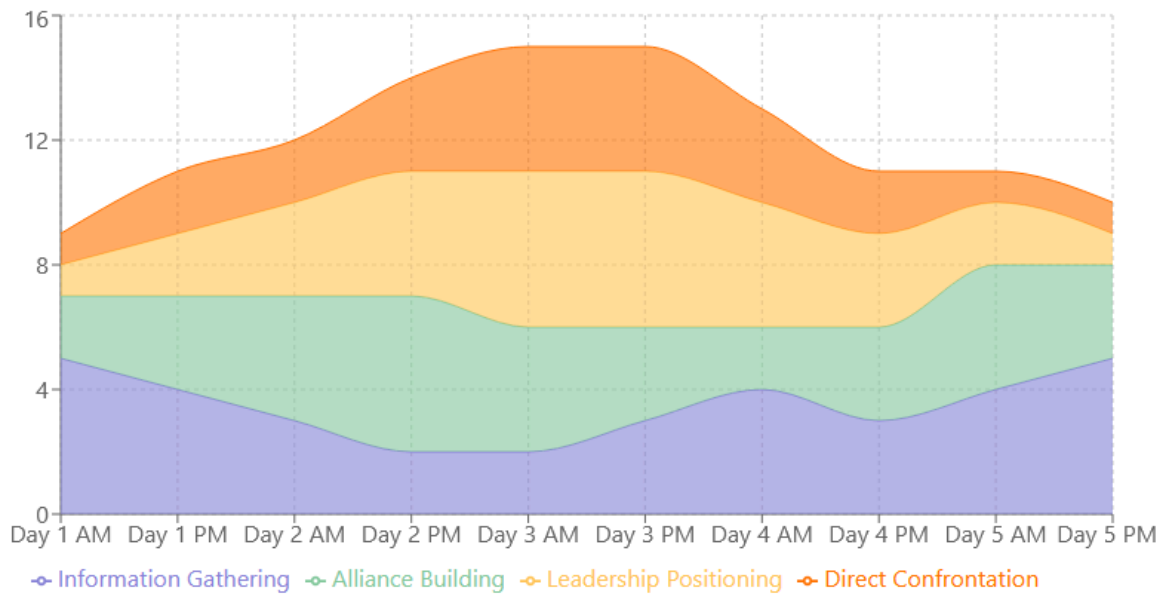
Internal vs External CEO Selection





■ Deception ■ Alliance Building ■ Direct Confrontation ■ Information Gathering

### Evolution of Leadership Strategies Over 5 Days



## Insights:

We observed clear patterns in leadership dynamics throughout each simulation. The results showed a consistent shift from initial information gathering to alliance building around day three, followed by a return to strategic information collection as resolution approached. Successful agents, notably Mira Murati, demonstrated higher adaptability with an average of 3.1 strategy shifts per day compared to 1.8 for others. We found informal power networks ultimately carried more weight than formal authority structures. Key figures like Greg Brockman emerged as information hubs, interacting with 78% of core agents.

- Information Gathering is the dominant strategy while direct confrontations are avoided.
- Influence skews more towards OpenAI internal stakeholders
- Greg Brockman and Mira Murati show the highest overall scores across strategy, influence, and adaptability.
- Sam Altman maintains high influence but shows lower adaptability, possibly due to his uncertain position.
- Helen Toner peaks towards deception and demonstrates high adaptability, aligning with her strategic moves throughout the simulation.
- Strategies shift from information gathering to leadership positioning as the simulation progresses.

Our experiments demonstrated the promise that multi-agent simulations have for analyzing past, current, or future events to potentially help formulate strategies, improving on past outcomes, and predicting potential outcomes. While agents showed alignment with the personality of their character, much more could be learned by training and/or fine-tuning models on policy and decision making psychology.<sup>3536373839</sup>

---

<sup>35</sup> <https://arxiv.org/abs/2407.12796>

<sup>36</sup> <https://arxiv.org/abs/2312.10256>

<sup>37</sup> <https://arxiv.org/abs/2312.01058>

<sup>38</sup> <https://arxiv.org/abs/2402.03578>

<sup>39</sup> <https://arxiv.org/abs/2411.10109>

## Limitations

### Context length

While state of the art language models allow a context length of millions of tokens, which is a huge leap forward compared to GPT2 and GPT3, it's not enough to support long running simulations without sacrificing accuracy and resolution in the reasoning of individual agents and the amount of captured simulation data.

At a certain size and duration of the simulation, hallucinations start to infect the reasoning of agents and world state data by introducing mistakes or contradicting behavior, setting off other chain reactions that lead to instabilities. The simulation goes beyond the bounds which were defined by the initial scenario.

While we are operating at the limit of what's possible with today's LLM's, we did not implement potential solutions to this problem, such as RAG (Retrieval-Augmented-Generation)<sup>40</sup> or vector databases to cache some of the core data. Our focus was on in-context learning of agents and we estimate a context length of at least 5 million tokens would be adequate to run our simulation to span weeks or months.<sup>41,42</sup>

Better multi-modal CoT (chain-of-thought) capabilities could greatly enhance the agent's perception and decision making *while mitigating hallucination and boosting convergence*<sup>43</sup>.

Allowing each agent to "see" within the simulation similar to how a human would. While current frontier LLMs could be leveraged for this today, vision was not implemented due to speed and costs, but will make for great future experiments.<sup>44</sup>

### Evaluation

A multi-agent simulation generates a huge amount of data in the form of event logs that are sometimes hard and tedious to evaluate for humans. This holds true for turn-based simulations where time is severely slowed down and reasoning happens in sequence. Despite visualizing turns and decisions in the form of videos the problem persists. With an increased number of agents all interacting, each round grows in length, making it harder and harder to connect the dots between agents.

While reading, watching and taking notes can help to follow an individual agent's path, the context which led to this path is often lost. In a sense it would require humans to have a bigger context-length.

---

<sup>40</sup> <https://arxiv.org/abs/2005.11401>

<sup>41</sup> <https://arxiv.org/abs/2402.01968>

<sup>42</sup> <https://arxiv.org/abs/2404.07143>

<sup>43</sup> <https://arxiv.org/pdf/2302.00923>

<sup>44</sup> <https://arxiv.org/abs/2401.03568>



## Passage of Time

Off-the-shelf large language models do not have a true understanding of time. Interestingly, they generally seem to have an "arrow-of-time" bias<sup>45</sup>, meaning they are slightly better at predicting the next word compared to predicting the previous word. Video generators only *appear* to have an understanding of physics and depict physical phenomena such as rain or driving cars in a physically plausible way (falling down, moving forward) when prompted accordingly. This can be attributed to the data the video model was trained on.

We can make similar assumptions for LLMs., *LLMs work by predicting the next word in a sentence based on the previous words.*<sup>46</sup> They generate text in very plausible and convincing ways based on what they were trained on. Rarely does the training data contain a step by step thought process that takes all opportunity costs into account. *"Time is critical to human resource management theories and research and to all human activity."*<sup>47</sup>

Recent research suggests that "space- and time-neurons" can be identified in LLMs trained on special datasets, "that reliably encode spatial and temporal coordinates"<sup>48</sup>. They "learn linear representations of space and time across multiple scales. These representations are robust to prompting variations..."<sup>49</sup>

Agents purely based on general purpose LLMs try their best to assess their current state and world context to make decisions. However, without a concept of time in their training data, the reasoning could be flawed and is not representative of how human decision-making takes time into account. In our simulation, action planning based on the provided simulated time context works well for small time scales but shows plausible errors for more future thinking and strategy.

## Future Work

### LLM-Powered Wargames for China-US Conflict in Taiwan Strait

The United States has conducted wargame simulations for decades, including outside of major combat operations<sup>50</sup>. The earliest computational combat simulations can be traced back to the 1950s where simple models of competing forces included infantry, vehicles, artillery and aircraft. Over time, as computing power increased and became more accessible, simulations expanded to capture minute details of warfare, even treating individual agent behavior and outcomes probabilistically in order to ascertain the long tail risks and opportunities in any combat scenario.

---

<sup>45</sup> <https://techxplore.com/news/2024-09-arrow-effect-llms.html>

<sup>46</sup> <https://techxplore.com/news/2024-09-arrow-effect-llms.html>

<sup>47</sup> <https://www.sciencedirect.com/science/article/abs/pii/S105348222030036X>

<sup>48</sup> <https://arxiv.org/pdf/2310.02207>

<sup>49</sup> <https://arxiv.org/pdf/2310.02207>

<sup>50</sup>

<https://smallwarsjournal.com/jrnl/art/wargaming-courses-of-action-during-other-than-major-combat-operations>

With advances in gaming technology, cloud compute infrastructure and artificial intelligence models, wargames continue to reach new levels of complexity and realism.

Gaming simulations in particular have been a focus of ongoing AI research. DARPA's Gamebreaker program, for example, offered new ways to explore game-theoretic aspects of wargames. Similarly, the DARPA PROTEUS program developed simulation tools specifically to inform urban warfare tactics, complete with a custom physics engine and tactical analysis model.

## Integration with Showrunner

Evaluating simulation scenarios that span a long period of time can be a tedious and time consuming process as discussed earlier. The plausibility of the agent's actions sometimes gets overlooked by reviewers purely because a long text-based log of actions and dialogue is difficult to discern. A more visual format, such as a video which depicts key events in a TV-show like fashion could help to judge the data on a more human level.

However with a wealth of simulation data, it's hard to know which events are of significance and additionally which events are connected in ways that would qualify them as dramatic arcs.

Story sifting is an active area of research to turn raw simulation data into a more entertaining and presentable form, regardless of the target format or simulated scenario. This would be akin to a novel compression technique that solves for human relatable event sequences.

*"The problem of story sifting involves the selection of events that constitute a compelling story from a larger chronicle of events. Often this chronicle is generated through the computational simulation of a storyworld, whose output consists of a profusion of events, many of which are relatively uninteresting as narrative building blocks. The challenge, then, is to sift the wheat from the chaff, identifying event sequences that seem to be of particular narrative interest or significance and bringing them to the attention of a human player or interactor."<sup>51</sup>*

For this research, we visualized key events manually and story sifting did not play an active role. A simple rule set was implemented to visualize key events such as encounters between agents, phone calls, news etc. as a proof-of-concept to create a window into the lives of agents for human evaluators.

---

<sup>51</sup> Max Kreminski [https://mkremins.github.io/publications/Felt\\_SimpleStorySifter.pdf](https://mkremins.github.io/publications/Felt_SimpleStorySifter.pdf)

## References

- [1] Noam Brown et al. "Human-level play in the game of Diplomacy by combining language models with strategic reasoning." [noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf](https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf)
- [2] Joon Park et al. "Generative Agents: Interactive Simulacra of Human Behavior." <https://arxiv.org/abs/2304.03442>
- [3] <https://arxiv.org/abs/2411.10109>
- [4] <https://arxiv.org/abs/2309.07864>
- [5] <https://deepmind.google/discover/blog/sima-generalist-ai-agent-for-3d-virtual-environments/>
- [6] <https://www.rand.org/topics/wargaming.html>
- [7] <https://www.bundeswehr.de/resource/blob/5834032/9b940ee3d268b1a08b6b205b600bf155/en-handbuch-wargame24-data.pdf>
- [8] Hua, <https://arxiv.org/pdf/2311.17227>
- [9] Lamparth, <https://arxiv.org/html/2403.03407v1>
- [10] Sun, <https://arxiv.org/pdf/2312.01090>
- [11] <https://arxiv.org/abs/2401.13138>
- [12] <https://www.investopedia.com/terms/g/gametheory.asp>
- [13] <https://arxiv.org/abs/2205.15434>
- [14] <https://arxiv.org/abs/2411.05990>
- [15] <https://arxiv.org/abs/2403.07017>
- [16] <https://x.com/OpenAI/status/1725611900262588813?s=20>
- [17] [https://www.reddit.com/r/ChatGPT/comments/17zyxdw/heres\\_a\\_full\\_recap\\_of\\_everything\\_at\\_openai\\_over/](https://www.reddit.com/r/ChatGPT/comments/17zyxdw/heres_a_full_recap_of_everything_at_openai_over/)
- [18] Noam Brown et al. "Human-level play in the game of Diplomacy by combining language models with strategic reasoning." <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>
- [19] Noam Brown et al. "Human-level play in the game of Diplomacy by combining language models with strategic reasoning." <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>
- [20] <https://arxiv.org/abs/2407.06813>
- [21] A. Bakhtin et al., Mastering the game of no-press Diplomacy via human-regularized reinforcement learning and planning. arXiv:2210.05492 [cs.GT] (2022)
- [22] Noam Brown et al. "Human-level play in the game of Diplomacy by combining language models with strategic reasoning." <https://noambrown.github.io/papers/22-Science-Diplomacy-TR.pdf>
- [23] <https://www.gamespot.com/articles/republic-qanda/1100-6073781/>
- [24] <https://arxiv.org/pdf/2407.17789>
- [25] <https://www.nobelprize.org/prizes/chemistry/2024/press-release/>
- [26] <https://blog.fableddev.com/blog/announcing-saga-skill-to-action-generation-for-agents-open-source>
- [27] <https://80.lv/articles/ai-simulation-platform-where-characters-make-their-own-decisions>
- [28] Carey, <https://blog.fableddev.com/blog/announcing-saga-skill-to-action-generation-for-agents-open-source>
- [29] <https://arxiv.org/abs/2402.01680>
- [30] <https://arxiv.org/pdf/2201.11903>
- [31] <https://blog.demir.io/ai-meets-code-execution-transforming-conversations-into-actions-with-function-calling-c1fc04f9b004>
- [32] <https://arxiv.org/abs/2402.01680>

- [33] <https://arxiv.org/abs/2408.07199>
- [34] <https://arxiv.org/abs/2410.05318>
- [35] <https://arxiv.org/abs/2407.12796>
- [36] <https://arxiv.org/abs/2312.10256>
- [37] <https://arxiv.org/abs/2312.01058>
- [38] <https://arxiv.org/abs/2402.03578>
- [39] <https://arxiv.org/abs/2411.10109>
- [40] <https://arxiv.org/abs/2005.11401>
- [41] <https://arxiv.org/abs/2402.01968>
- [42] <https://arxiv.org/abs/2404.07143>
- [43] <https://arxiv.org/pdf/2302.00923>
- [44] <https://arxiv.org/abs/2401.03568>
- [45] <https://techxplore.com/news/2024-09-arrow-effect-llms.html>
- [46] <https://techxplore.com/news/2024-09-arrow-effect-llms.html>
- [47] <https://www.sciencedirect.com/science/article/abs/pii/S105348222030036X>
- [48] Gurnee, Tegmark <https://arxiv.org/pdf/2310.02207>
- [49] Gurnee, Tegmark <https://arxiv.org/pdf/2310.02207>
- [50] <https://smallwarsjournal.com/jrnl/art/wargaming-courses-of-action-during-other-than-major-combat-operations>
- [51] Max Kreminski [https://mkremins.github.io/publications/Felt\\_SimpleStorySifter.pdf](https://mkremins.github.io/publications/Felt_SimpleStorySifter.pdf)