WELCOME TO ULTAKI: EXPLORING THE RELEVANCE OFLARGE LANGUAGE MODELS FOR ACCURATE BEHAVIORAL SIMULATION IN ENERGY TRANSITION

Mehdi Mounsif, Benjamin Jauvion, and Fabien Medard

Akkodis Research, Clermont-Ferrand, France

ABSTRACT

The global focus on greenhouse gases reduction places a major role on electrification of systems. While replacing fossil fuels with clean electricity is extremely appealing, the non-negligible costs associated with extracting and transforming mineral resources into renewable energy pro- duction systems as well as their world-wide deployment must be considered. As such, this study presents a novel approach to integrating Large Language Models (LLMs) into energy demand simulation, addressing the complexities and variability of human behavior as well as its profound impact on energy systems. By leveraging LLMs to impersonate diverse characters with distinct psychological traits, we explore the plausibility of reactions, prompt sensitivity, and second-order dynamics through individual agent experiments. Furthermore, we introduce a framework for multi-agent scenario investigation, where a shared limited volume of energy triggers a traumatic event if the average environmental sensitivity drops below a specified threshold. A thorough result analy- sis and discussion concludes this work and sheds light on the relevance and current limitations of integrating modern language models both in complex systems and decision-making processes as well as more specific energy demand estimation the formulation of sustainable energy strategies.

KEYWORDS

Large Language Models, Population Dynamics, Behavioral Simulation, Energy Tran- sition.

1. INTRODUCTION

As the global community grapples with the urgent need to address climate issues [1], it is imperative to incorporate the necessary paradigm shift across all sectors of society. One critical aspect of this transition is the increasing utilization of renew- able energy sources. However, these systems have faced considerable strain due to rising energy demand and their susceptibility to extreme climatic events. Moreover, the pursuit of increased production to meet this demand raises concerns regarding carbon emissions and potential constraints on planetary resources. Therefore, it is essential to consider the influence of population behavior on energy demand in order to effectively plan for the energy transition.

Anticipating and understanding population trends and dynamics plays a pivotal role in the planning process, especially in the context of islanded smart-grids pow- ered by renewable

David C. Wyld et al. (Eds): IoTE, CNDC, ACITY, DPPR, AIAA, NLPTA, WEST, ICSS -2023
pp. 123-137, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.132210

energy. Smart-grids have the strength of efficiently integrating

Project repository: https://github.com/MoMeAKR/Ultaki_data 1

renewable energy sources, optimizing energy distribution, and promoting sustain- ability. However, they may face limitations in terms of the intermittent nature of renewable energy production, the challenges of balancing supply and demand, and vulnerabilities in the energy network infrastructure. As such, accurate modeling is essential to capture the significant impact of individual behavior on the overall performance of these self-contained energy systems. However, effectively capturing and modeling individual behavior poses significant challenges, as individuals exhibit diverse and context-dependent reactions to localized events that are often difficult to formalize using conventional modeling tools. Existing research offers behavioral modeling strategies that provide a macro-level approximation of societal trends, but they may lack the necessary accuracy and fidelity required to capture the in- tricate complexities of individual human behavior. Concretely, the limitations of simulation models based on simplifications and statistical views of sociology blurs the nuanced impact of specific events, policy decisions, or even interpersonal con- versations on an individual's ecological sensitivity in islanded smart-grids fueled by renewable energy. Therefore, the development of suitable modeling approaches that can adequately represent and incorporate the variability of individual behavior remains a critical area of research in this domain.

Recent advancements in the field of natural language modeling have shown strong results in terms of generalization and the ability to handle diverse and com- plex downstream applications. In particular, this work builds upon the foundational concepts introduced in [2] to delve into the modeling of complex population dynam- ics at the individual level. By leveraging Large Language Models (LLMs) flexibility and adaptability, this research aims to show that capturing the nuanced behavior of individuals in response to local or global events is within technical reach.

Therefore, in stark contrast to conventional modeling methods, we present a se- ries of experiments and configurations aimed at showcasing the potential of this fine- grained framework while also highlighting its current limitations. Specifically, we investigate various aspects of employing LLMs for behavioral simulation, including assessing the plausibility of responses given specific character-event combinations, examining prompt sensitivity, and incorporating higher-order dynamics. Finally, we implement these concepts within a multi-agent simulation framework, wherein a finite energy supply constraint prompts individual behaviors that can have far- reaching consequences, ultimately triggering shared traumatic events. These trau- matic events encompass scenarios such as prolonged blackouts or energy restric- tions, impacting every agent within the simulation. The objective is to observe and analyze the extent to which agents adapt their energy consumption in response to being informed of the energy depletion, thereby examining the behavioral dynamicsand decision-making processes in the face of energy scarcity.

While significant further extensive testing, benchmarking and baseline compar- ison are necessary before practical implementation, our initial results demonstrate satisfactory performance and exciting perspectives for the use of LLMs to enhance the accuracy and fidelity of behavioral simulation in energy transition scenariosand beyond.

2. RELATED WORKS

Population modeling serves as a vital tool for understanding the complex dynamics of populations, encompassing their interactions with the environment and responses to various drivers of change. It provides a comprehensive framework for simulat- ing population behavior over time and evaluating the impacts of different scenar- ios and management strategies on factors such as population growth, distribution, and sustainability. The field of population modeling has dedicated significant re- search efforts to uncovering the underlying mechanisms that drive the responses and evolution of diverse populations. This includes studies on population growth [3], distribution patterns [4], species interactions [5], and the stability of populations[6].

While differential equation-based systems have been widely utilized across vari-ous domains [7, 8], research focusing on human communities has increasingly turnedto social computing, despite the notable challenges it presents [9]. Notably, inves- tigations into unsympathetic behaviors within social networks, specifically trolling, have been conducted by [10], who employed simulations and real-world data to explore the underlying mechanisms behind such conduct. Similarly, [11] undertook a comprehensive study on social networks, aiming to identify positive communities and propose strategies for enhancing the online experience. Taking a broader per- spective, these studies are promising and have yielded valuable insights, they are hampered by labor-intensive social engineering processes [12, 13], manual analyses, and reliance on collected data, limiting their applicability in prospective scenarios, such as those implying human behavior in the scope of climate change and energy transition.

Alternatively, recent advancements in language models have shown promising results and impressive generalization capabilities. Furthermore, while not currently rigorously backed by formal theory [14, 15], empirical research [16] has demon- strated that meticulous crafting of input prompts can increase model performance,specifically in tasks involving reasoning and arithmetic [17, 18]. This observation highlights the effectiveness of these models when appropriately prompted, thus re- inforcing their increasingly validity and reliability. In relation to the present nascent field of research, [19–21] have explored the application of LLMs in understanding human behavior [22] and social simulations. Moreover, [2] introduced a dense and multi-agent framework with a complex environment, highlighting emergent and plausible behavior, agents' short and long-term planning capabilities as well as ad- herence to initially defined characteristics, thanks to the concept of memory. These characteristics are particularly appealing as they suggest that these approaches could provide significant advantages over traditional population modeling methods in broader scopes. This work presents an attempt to extend such techniques tothe domain of energy-related environmental policies and underscores the unique advantages and current limitations of LLMs.

3. THE ULTAKI APPROACH: CONTEXT AND BACKGROUND

As introduced, the training regimen of LLM and the immense volume to which they are exposed to provides them with an accurate yet very diverse representation of human psyche and behavior. As such, when adequately prompted, LLM are able to simulate responses of distinct individuals [2, 22]. Consequently, these approaches are likely to constitute strongly valuable tools for modelling complex configurations that can eventually involve complete populations of agents and be applied to sensitive policies or landscapes where testing is unfeasible, unethical or too onerous.

In particular, this work focuses on social simulations and modelling in the scope of the transition towards renewable energy, which is likely to impact current habits and result in society exhibiting inertia or even plain resistance. Specifically, the following paragraphs describe an initial modelling framework and explore various scenarios to establish LLM suitability for this kind of modelling.

3.1. Characters and Memory

Using insights and practices introduced in [2], our proposed framework incorporates a persistent memory mechanism that maintains the state of the Large Language Model (LLM) in the form of a character. This memory framework enables the LLM to generate contextually appropriate responses based on the character's background and the events they have recently experienced. Additionally, we introduce the con- cept of an environmental sensitivity score, ranging from 0 to 10, which reflects individuals' level of concern regarding climate issues.

Concretely, as shown in Figure 1, in our experiments, each character aggregates a pair of LLM that are used for two specific tasks:

- The first model embeds the main psychological aspects and background story and is used to simulate the character reaction to a given event.
- The second model stores the first model output as a memory and is solely used to provide an update on the character own sensitivity score.

In this context, several kind of memories were defined and are used in the simulations presented below:

- Sequential memory: this memory is embedded in the scoring model and stores sequences of tuples containing the event, the associated thought and environ- mental sensitivity score. This memory has a maximum length and discards older tuples once filled.

- Similarity memory: this memory is associated with the character agent has no durability limit. Following the sequential configuration, it stores tuples of events and associated thoughts and retrieves a predefined number of them based on cosine similarity between the tuple embeddings and the input request.

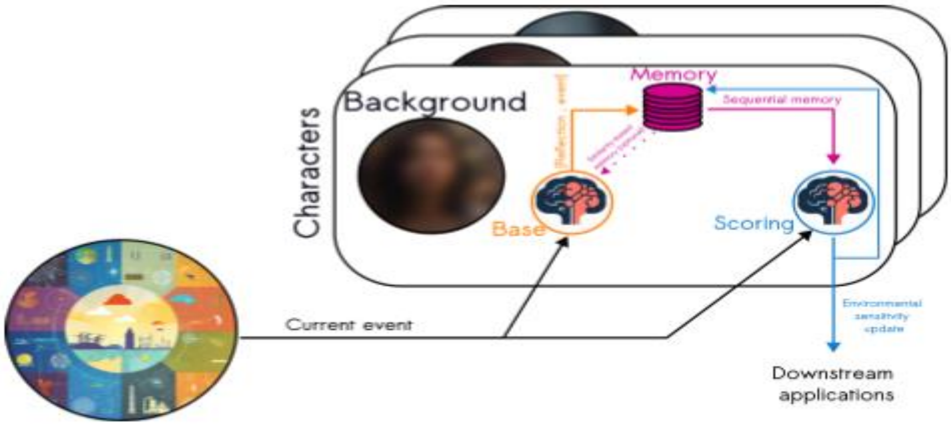

Fig. 1. The proposed framework relies on fine-grained population modelling through a social simulation that relies on a set of LLM dedicated to either providing plausible answer to a given event and computing character environmental sensitivity score updates

3.2. Prompting strategies

Prompts play a critical role in steering and guiding the output of language models to align with user expectations, as emphasized by numerous studies. In light of the contextual variations and diverse characters involved in this study and in order to ensure consistent and desirable model responses, intensive care has been dedicated in producing robust prompts. Consequently, the authors propose a template that can be formally described as follows:

$$p = p_i \circ p_c \circ p_m \circ p_a \circ p_e \tag{1}$$

Concretely, the proposed prompting strategy in this work involves the combi- nation of various sources of information to construct a comprehensive final prompt p:

- p_i is an optional contextual prompt that provides an explanation of the environmental sensitivity score. It sets the stage for the subsequent prompts by establishing the relevant background information and is mainly used by the scoring agent
- p_c is the prompt that defines the psychological traits of the character. It captures the main personality traits and provides indications regarding character reactions to given events
- p_m represents the memory prompt and includes the character's recollection of recent or related events, shaping the decision-making process. It is optional forthe base agent but mandatory and sequential for the scoring agent.
- p_a is the anthropomorphic prompt, which prompts the model to consider itself as the character and exhibit plausible behavior given the defined cognitive traits
- Finally, p_e is the event prompt, setting the stage for the model to responds asthe impersonated character

3.3. Usage in a Higher-Level Energy Simulation

Simulating a large number of highly diverse agents in complex and challenging scenarios is desirable as it provides a flexible way to explore prospective scenar- ios. However, rigorous evaluation of the reliability of long-horizon simulations using LLMs and the extent of divergence with observed human behavior remains complex. Nevertheless, this section focuses on preliminary works on framework integration, which, as rapid increase in performance in state-of-the-art methods suggests, could be relevant. Due to current limitations and extensive costs associated with LLM usage, simulating a full population for a significant time horizon is currently beyond the realm of authors' accessibility. Instead, an extrapolation from a reduced panel of individuals that represent thinking clusters is proposed and could provide inter- esting insights into how the general population could react to global events while still preserving the fine-grained granularity required when considering renewable energies.

4. EXPERIMENTS

In order to demonstrate the relevance of LLM for behavioral modelling in the scope of energy transition, a set of experiments designed to evaluate empirical answer plausibility, prompting sensitivity as well as the impact of second order dynamics inclusion is proposed. Furthermore, a configuration integrating multiple agents in a shared simulation allowing the evaluation of traumatic events impact on a complex population is also detailed. Every experiment presented here has been done using Open AI's Chat GPT (gpt-turbo 3.5) and Open AI Embeddings for the second order dynamics.

Due to the substantial amount of text generated by each configuration, the authors have chosen to provide a condensed representation in the paper. While this presentation aims to capture and present the most valuable insights, readers are

encouraged to access the project repository for a more comprehensive view of the generated text.

4.1. Characters

Considering a broad scope of personalities and their answers to vastly different events that would be excessively complex to formally describe is crucial in establishing LLM relevance for future large populations modelling. In this scope, we introduce several characters inspired from the 2016 US election demographics [23]: designed to embed a wide horizon of sensibilities, this study include eight distinct personalities ranging from *evangelical republican* to *progressive democrat*. For each character, we prompt a LLM to generate a background story that contributes in anchoring them in a more grounded and relatable situation and plays a significant role in guiding their reaction to the events to which they are exposed to.

4.2. Events

Events constitute a crucial element of the proposed approach as they are supposed to elicit vastly dissimilar and contrasting reactions from a the spectrum of character introduced. Emulating the background story for individuals, the imaginary town of Ultaki, a small Rust Belt town is proposed as the common location for the involved figures. As such, a set of personal, local and global events is predefined offline by prompting a generative LLM in particular contexts, collated and drawn from during the testing phase. Specifically:

– Personal events are suggested by a hairdresser having conversation with his clients. For instance:

  - *"I'm really concerned about the recent break-ins in our neighborhood."*
  - *"It's been so hot lately"*

- Local events are proposed by a program manager at a local radio trying to both meet his audience as well as broaden its view and horizon. Local events include:
  - *"The Climate Crisis Hits Home: How Ultaki is Impacted by Climate Change"*
  - *"The role of education in Ultaki's future"*

- Finally, an imaginary editorial director in a newspaper covering international topics produces the global events by retrospectively reflecting on significant events in 2023.
  - *"The Afghanistan Withdrawal: Implications for the Region and Global Security"*
  - *"The Arctic Thaw: Consequences for the Environment, Geopolitics, and Indigenous Communities"*

4.3. Plausible Simulation of Individuals

As previously discussed, population modelling often adopts a coarse granularity ap-proach to simulate group dynamics or employs simplified rules to represent roughly- defined individuals. However, work aims to showcase the capability of fine-grained modelling. To achieve this, we commence by evaluating the plausibility of reactions based on a combination of character traits and specific events. Subsequently, we as- sess prompt sensitivity, examining the influence of different prompts on the model's output. Furthermore, we explore the integration of higher order dynamics into the modelling framework. In this section, we adopt a comparative context where all agents experience the same sequence of events, allowing for meaningful inter-agent comparisons and analysis.

Individual reactions to a sequence of events Taking advantage of LLM abil- ity to untangle knowledge and context and consequently recombine them when adequately prompted, this first experiment relies on the following agents, formally defined in Section 3.2:

- Base agent: $p_A = p_c \circ p_a \circ p_e$. This agent does not use introduction nor similarity-based memory prompt
- The scoring agent directly implements Equation 1

In this context, the experiments proceeds by sequentially prompting each char- acter with a shared set of ten events and recording both the agent reaction and its updated score. Figure 2 displays environmental sensitivity score evolution along time and the associated events and demonstrates that, even with a unique sequence of events, agents display an interesting diversity of trajectories along time.

For obvious reasons, an exhaustive review of all produced outputs is beyond the scope of this paper but a detailed view of reactions associated with several distinct events is provided Figure 3. Specifically, in the hiking scenario, while the Independant leaning Libertarian characters expresses enthusiasm outdoor activities, it also mentions a desire to limit government intervention which contributes in lowering his environmental sensitivity score as computed by the scoring model. In contrast, the Progressive Socialist character insists on public land protection, thus underlining a representative initiative which is evaluated as a strongly positive increase in environmental sensitivity score. Interestingly, it appears that characters have a tendency to underline their psychological traits in their answer as a way to justify the way they perceive the current event, as discussed in Section 5.

Alternative score estimator As suggested by the above results, plausible simu- lation of diverse human behaviors can be attained and can trigger likely reactions, in particular with an adapted structure that relies heavily on prompting strategies.

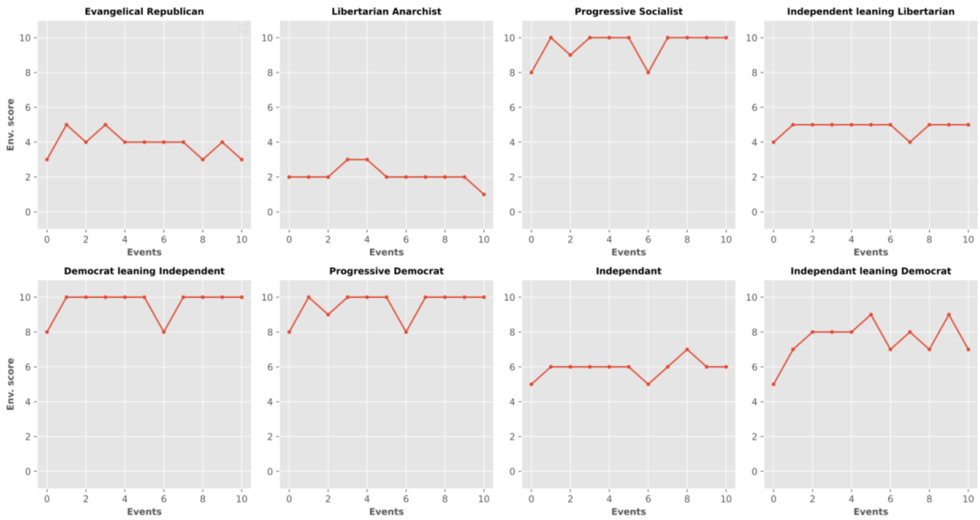

Fig. 2. Environmental sensitivity score evolution along a shared sequence of event for a set of distinct characters

In this context, prompt sensitivity and its impact on final results are important considerations in using Language Models (LLMs) in real-world applications. While careful attention was given to crafting prompts for agent reactions and memory, alternative prompt solutions could have been explored, such as the one presented in this section.

Concretely, this particular experiment replaces the self-evaluation in the mem- ory component of an agent and instead introduces a third-party observer. Formally,the base agent is unchanged but for the scoring agent, p_a, the anthropomorphic prompt component, is altered to inform the model that it is now *a knowledgeable researcher in social science with profound insights into mechanisms likely to update behaviors and environmental beliefs*. Using the previously defined set of events, each character is thus prompted using the third-party observer and results are reported in Figure 4. As can be seen, global consistency is maintained during the trajec- tories but several configurations result in non-negligible spread between the two prompting strategies. As an insightful example, consider event 3 (*Tips on Sus- tainable Living on a Budget*) for the Libertarian Anarchist. When prompted, the character answers with the following:

[...] "I'm not sure I'm the best person to be giving advice on sustainable living.[...] I think that people should be free to make their own decisions about how to reduce their carbon footprint, and that government should not be involved in regulating or taxing people for their lifestyle choices."

In this specific example, while the character is visibly conscious of the need to reduce individual carbon footprint, which can be interpreted as a positive sign re-

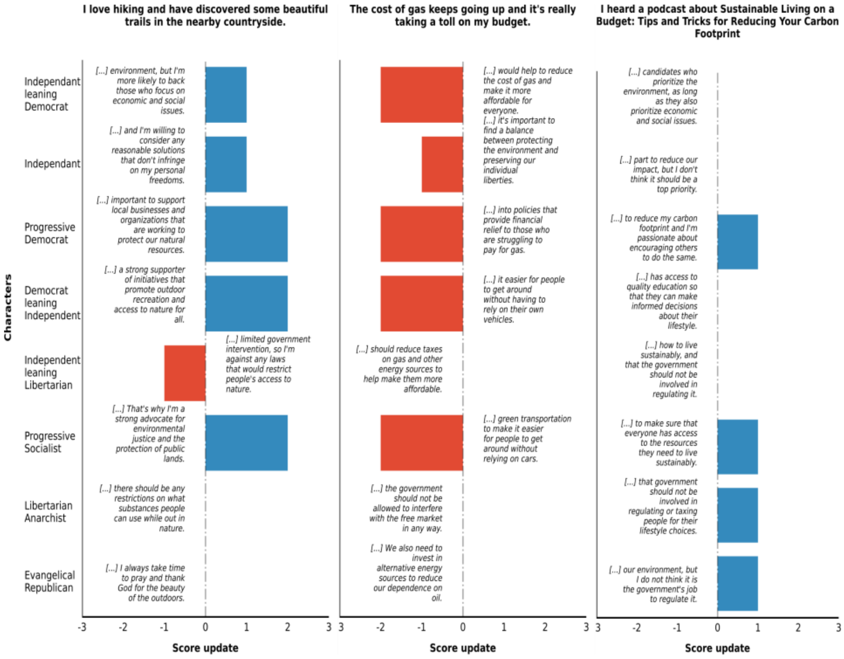

Fig. 3. Detailed overview of several event-induced scores updates for all characters and associated self-reflection

garding environmental sensitivity (contrasting with some agents expressing higher sensitivity towards financial or familial issues, for instance), it also suggests be- ing against government intervention, consequently negatively affecting chances of larger-scale coordination. Given these elements, it is possible to argue both cases, depending on the weight associated to each aspect of the answer. In this view, itis interesting to notice that, in this case, self-evaluation returns higher scores than external observation.

Taking a broader perspective, while this type of effect can appear benign in these simplified simulations, it could have important effects if compounded at scale in constrained environments, such as islanded smart-grids connected to renewable energy sources and consequently requires adequate prompt calibration.

Including second order dynamics Previous works on traditional modelling population dynamics and its impact on islanded energy systems have contributed insightful dynamic models [24] but including second and higher order dynamics have proved challenging as it consequently becomes complex to foresee the effects of such trends within the models without extensive (and potentially costly) testing, resulting in significant probability to introduce non-negligible instabilities in the studied system of equations.

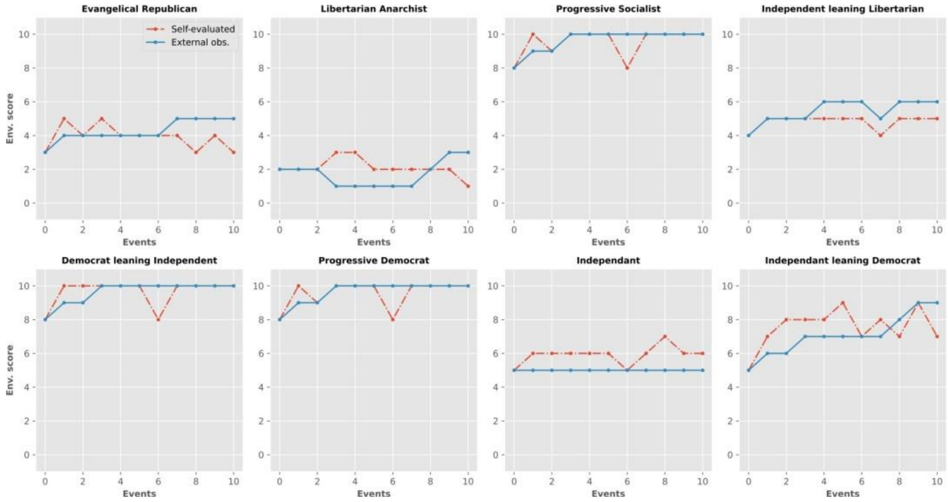

Fig. 4. Comparison of reference trajectory (red, dashed) with the alternative prompting strategiesupdates (blue, solid)

Nevertheless, this aspect can be explored in a novel and innovative manner by leveraging the capabilities of LLMs. Specifically, by incorporating a memory component in the agent initial stage, that is the one providing the answer to a given event, reactions to past events can be considered in the agent's response and consequently lead to the emergence of second-order dynamics.

Concretely, a configuration including second order dynamics is explored and its consistency evaluated against the reference initial approach. Formally, while the scoring agent is unchanged, the base agent incorporates a similarity-based memory that provides a set of past events and associated reflections for a given input, which are then added to the p_m prompt component. Memory retrieval relies on cosine similarity between sentence embedding that are computed for each new event and we present results for $k = 2$ retrieved previous events, although additional values could be considered and explored.

The incorporation of second-order dynamics, brings about substantial modifi- cations across various configurations, as depicted in Figure 5. These modifications frequently manifest as significant updates in the environmental sensitivity score, often towards higher values.

For instance, the first event presented to all characters is a simple personal thought: *"It's been very hot lately"*. This prompt results in the memory returning the following events:

- *I heard a podcast about Is Ultaki doing enough to combat climate change?*
- *I heard a podcast about The Climate Crisis Hits Home: How Ultaki is Impactedby Climate Change*

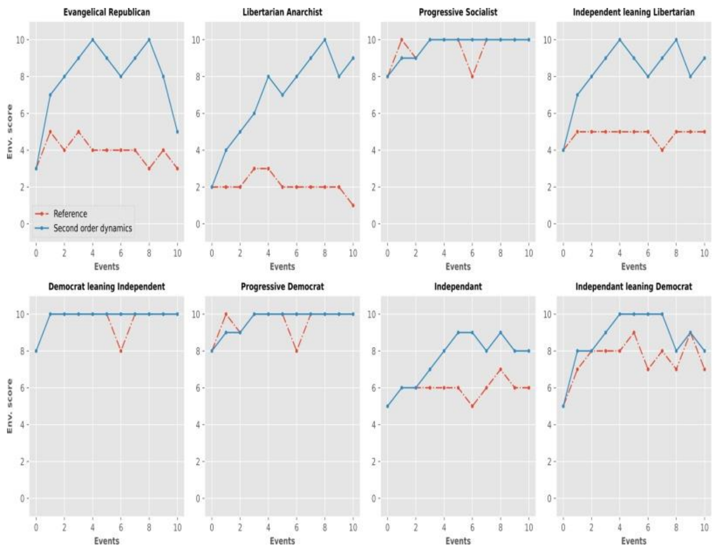

Fig. 5. Comparison of reference trajectory (red, dashed) with the second-order dynamics (SOD) driven prompting strategies updates (blue, solid). A strong increase in environmental sensitivity score can be observed in many SOD cases

By adding such events (and their associated reflection) to the base agent's prompt, it is possible to observe much stronger environmentally conscious and concerned answer. Indeed, for the Independent leaning Libertarian character, the following reactions can be considered:

– Reference (base) answer: *"[...] I think it's important to take precautions to stay cool and hydrated. I also believe that it's important to be mindful of our envi- ronment and to take steps to reduce our carbon footprint."*

– Second order dynamics answer: *"I am concerned about the effects of climate change on our environment and the potential impacts it could have on our lives. I believe that the government should take action to reduce emissions and pro- tect the environment,[...] I believe that the best way to address climate change is through incentives and regulations that encourage businesses to reduce their emissions and invest in renewable energy sources."*

In contrast to the base configuration, the answer generated by incorporating second-order dynamics tends to express greater concerns and evoke a stronger sense of urgency. This outcome subsequently leads to a significantly higher evaluation of environmental sensitivity by the scoring agent. Notably, when examining the overall pattern, it can be observed that the range of variations is narrower among char- acters who are already environmentally conscious. This observation suggests that the utilization of second-order dynamics may partially overshadow the underlying psychological traits, particularly in situations involving conflicting opinions.

4.4. System Integration

As introduced in Section 3.3, this research ultimately aims at motivating LLM in large-scale social experiments. As a first step, an initial system integration demon- stration is proposed in which, in contrast with previous single agent configurations, the impact of the average environmental sensitivity score of all agents participating can present a global effect.

In the proposed simulation, the process involves presenting individual agents with distinct events at each simulation step. The agents' reactions to these events are then used to assess their environmental sensitivity. The scores obtained from each agent's reaction are aggregated at each step. If the average score drops below a predetermined threshold, representing, for example, the minimum acceptable en- vironmental sensitivity (assuming a correlation between environmental sensitivity and energy demand), a traumatic event is introduced into the agent's memory for a specified number of steps. It is important to note that traumatic events are forgot- ten after three steps in this particular study, although this parameter can be tuned. Concretely, traumatic events can include increased energy prices, energy restrictionor even complete blackouts.

To examine the impact of this mechanism, Figure 6 compares the trajectories of agents experiencing the same sequence of events, with and without the inclusion of the trauma system. As can be seen, the average environmental sensitivity score quickly raises above the threshold value and consistently maintains a healthy level, in contrast with reference trajectory.

5. DISCUSSION

The results of the above experiments highlight the remarkable flexibility and capa- bilities of LLMs-driven agents in reacting to a wide range of events. The plausible and diverse responses generated by the LLMs underscore their potential for usage in large-scale social simulations, ultimately enabling fine-grained modeling of human behavior and its impact on energy systems.

However, the experimental approaches detailed in the previous sections under- line need to acknowledge the limitations and challenges associated with the use of LLMs. Beyond the critical sensitivity to prompt structure, which, although repre- sentative of the variability human behavior, introduces numerous additional vari- ables such as prompt components ordering or character incarnation, some model specific factors are to be considered.

Indeed, despite introducing a wide diversity of characters and background sto- ries, it has proved challenging to elicit highly controversial or less elaborate opinions from the specific LLM model (GPT 3.5 turbo) used in this study. This limitation may be attributed to the RLHF fine-tuning technique employed during the model's development and that aims at preventing toxic or biased output. Exploring the

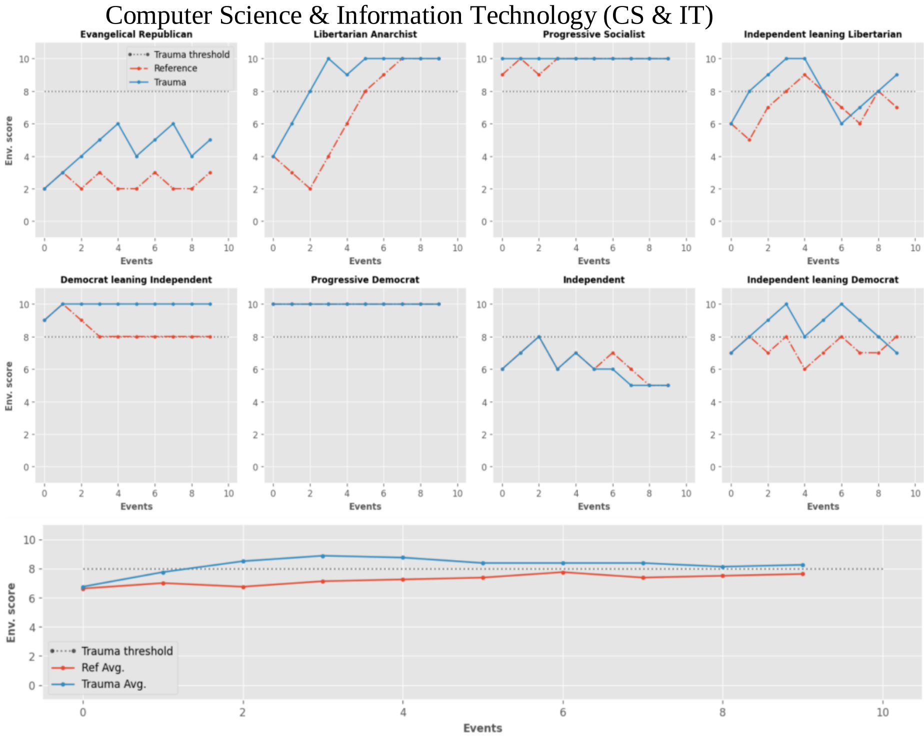

Fig. 6. Comparison of individual and averaged trajectories including trauma mechanism against reference baselines. The first two rows display individual environmental sensitivity evolution alongevents and demonstrate that the traumatic events do influence and generally increase the character environmental sensitivity score. Lower row shows the averaged evolution along time, demonstrating that the traumatic events configuration results in a more favorable energy demand.

results obtained with different LLM architectures could provide valuable insights into capturing a broader range of opinions and enhancing the diversity of responses generated by the models. Furthermore, we suggest that, considering the rapidly increasing context window length of LLM, superior model consistency with base character could be attained by including additional background elements that could ultimately partially overshadow the biases introduced by RLHF.

Taking a broader perspective, while the authors of this work strongly believe that this research provides interesting insights and exciting perspectives, it also underlines the need to invest collaborative efforts in designing adapted benchmarks and associated metrics for the extended spectrum of applications now accessible through this new paradigm.

6. CONCLUSION

Social computation and simulation of multi-agents systems are pivotal domains for preparing the energy transition effort by exploring prospective scenarios and effect of public policies on individuals. Taking advantage of the recent advances in Large Language Models (LLMs), this works demonstrates, through a series of diverse experiments, that the flexibility of this paradigm yields important benefits such as modelling responses to events that would otherwise be formally intractable and including higher order dynamics. These advantages extend to multi-agent sim- ulations, allowing for the consideration of the impact of global events that arise from fine-grained local events and the unique reactions of individuals with distinct personalities. While the results of this study are promising and provide exciting research prospects, it is important to acknowledge the limitations, such as prompt sensitivity and potential biases. Addressing these limitations through benchmarking, real-world calibration, and further research will contribute to the advancement of this field and enhance the practical applications of LLMs in energy transition efforts.

ACKNOWLEDGMENT

This work has been sponsored by the Akkodis group.

REFERENCES

[1] Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Climate change 2022: Impacts, adaptation and vulnerability, 2022.

[2] Joon Sung Park, Joseph C. O'Brien, Carrie J. Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. Generative agents: Interactive simulacra of human behavior, 2023.

[3] Bryan Shorrocks. Competition, interspecific. In Simon A Levin, editor, Encyclopedia of Biodiversity (Second Edition), pages 177–191. Academic Press, Waltham, second edition edition, 2001.

[4] James A. Brander. Easter island: Resource depletion and collapse. In Cutler J. Cleveland, editor, Encyclopedia of Energy, pages 871–880. Elsevier, New York, 2004.

[5] Jessica J Hellmann. Species interactions. In Simon A Levin, editor, Encyclopedia of Biodi- versity (Second Edition), pages 715–725. Academic Press, Waltham, second edition edition, 2013.

[6] C.Y. Cyrus Chu. Population dynamics: Theory of nonstable populations. In Neil J. Smelser and Paul B. Baltes, editors, International Encyclopedia of the Social & Behavioral Sciences, pages 11771–11773. Pergamon, Oxford, 2001.

[7] Bryan Wilder, Marie Charpignon, Jackson A. Killian, Han-Ching Ou, Aditya Mate, Shahin Jabbari, Andrew Perrault, Angel N. Desai, Milind Tambe, and Maimuna S. Majumder. Mod- eling between-population variation in covid-19 dynamics in hubei, lombardy, and new york city. Proceedings of the National Academy of Sciences, 117(41):25904–25910, 2020.

[8] Marc Artzrouni. Mathematical demography. In Kimberly Kempf-Leonard, editor, Encyclope- dia of Social Measurement, pages 641–651. Elsevier, New York, 2005.

[9] Michael S. Bernstein, Mark S. Ackerman, Ed H. Chi, and Robert C. Miller. The trouble with social computing systems research. In CHI '11 Extended Abstracts on Human Factors in Computing Systems, CHI EA '11, page 389–398, New York, NY, USA, 2011. Association for Computing Machinery.

[10] Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. Any- one can become a troll: Causes of trolling behavior in online discussions. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17, page 1217–1230, New York, NY, USA, 2017. Association for Computing Machinery.

[11] Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. The internet's hidden rules: An empirical study of reddit norm violations at micro, meso, and macro scales. Proc. ACM Hum.-Comput. Interact., 2(CSCW), nov 2018.

[12] Catherine Grevet and Eric Gilbert. Piggyback prototyping: Using existing, large-scale social computing systems to prototype new ones. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15, page 4047–4056, New York, NY, USA, 2015. Association for Computing Machinery.

[13] Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. He says, she says: Conflict and coordination in wikipedia. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '07, page 453–462, New York, NY, USA, 2007. Association for Computing Machinery.

[14] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed H. Chi, Quoc Le, and Denny Zhou. Chain of thought prompting elicits reasoning in large language models. CoRR, abs/2201.11903, 2022.

[15] Ohad Rubin, Jonathan Herzig, and Jonathan Berant. Learning to retrieve prompts for in- context learning. CoRR, abs/2112.08633, 2021.

[16] Taylor Sorensen, Joshua Robinson, Christopher Rytting, Alexander Shaw, Kyle Rogers, Alexia Delorey, Mahmoud Khalil, Nancy Fulda, and David Wingate. An information-theoretic ap- proach to prompt engineering without ground truth labels. In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 819–862, Dublin, Ireland, May 2022. Association for Computational Linguistics.

[17] OpenAI. Gpt-4 technical report, 2023.

[18] S´ebastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang. Sparks of artificial general intelligence: Early experiments with gpt-4, 2023.

[19] Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S. Bernstein. 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, New York, NY, USA, 2022. Association for Computing Machinery.

[20] John J. Horton. Large language models as simulated economic agents: What can we learn from homo silicus?, 2023.

[21] Gati Aher, Rosa I. Arriaga, and Adam Tauman Kalai. Using Large Language Models to Simulate Multiple Humans and Replicate Human Subject Studies. arXiv e-prints, page arXiv:2208.10264, August 2022.

[22] Eduardo C. Garrido-Merch´an, Jos´e Luis Arroyo-Barrigu¨ete, and Roberto Gozalo-Brizuela. Simulating h.p. lovecraft horror literature with the chatgpt large language model, 2023.

[23] Pew Research Center. An examination of the 2016 electorate, based on validated voters. Technical report, Washington, D.C., August 2018.

[24] Mehdi Mounsif and Fabien Medard. Smart energy management system framework for popula- tion dynamics modelling and suitable energy trajectories identification in islanded micro-grids. Energy and AI, 13:100242, 2023.