How do Abstraction and Emotions Travel Different Spaces?

Jiayi Chen
Economics
Grinnell College
Grinnell, USA
email: jc12807@stern.nyu.edu

Ella Berman
Computer Science
Grinnell College
Grinnell, USA
email: bermanel@grinnell.edu

Zi Ye
Computer Science
Grinnell College
Grinnell, USA
email: yezi@grinnell.edu

Kailee Shermak
Sociology
Grinnell College
Grinnell, USA
email: shermakk@grinnell.edu

Mahiro Noda
Computer Science
Grinnell College
Grinnell, USA
email: nodamahi@grinnell.edu

David Rothfusz
Computer Science
Grinnell College
Grinnell, USA
email: rothfusz2@grinnell.edu

Fernanda Eliott
Computer Science
Grinnell College
Grinnell, USA
email: eliottfe@grinnell.edu

***Abstract*—Emotions *per se* are an intriguing multidisciplinary topic; still, the combination of emotions and technologies brings more layers into the subject. For example, on the one hand, we may consider a socially mediated communication of emotions, and on the other, Artificial Intelligence (AI) systems quickly merging into our realities and spaces. Here, we design an approach aiming to spotlight networked emotions and emotion-driven AI systems. This research started from our goal to encode abstract and emotion-rich contexts into an AI agent modeled after emotions and designed to cooperate purposefully with humans. Here, we narrow our questions to "How do abstraction and emotions travel different spaces?". We present our five-phase project idea to investigate that question, which explores distinct spaces: images, textual descriptions, 3D scenes, and mental models.**

***Keywords-abstraction; emotions; mental models; 3D-scenes*.**

I. Introduction and Motivation

When telling a story, we do not need to provide every single detail; we expect others to fill in the gaps and evoke mental models consistent with the story. For example, if it involves a library, it may be associated with a quiet place filled with books and other associated behaviors/rules. Mental models are "internal representations of the external world consisting of causal beliefs that help individuals deduce what will happen in a particular situation" [1]. We use 'mental models' as an umbrella term that covers spatial mental models and mental representations of environments or 'cognitive collages' [2]. Meanwhile, emotional mental models cover emotions and feelings connected to mental models: "Mental models cause certain expectations/thoughts of how things should look like/work and connect certain emotions with this. Consequently, a mental model is a cognitive and an emotional framework in the brain, influenced by person's personality (genes) and the environment including social variables" [3].

However, what if we wanted to design an AI system capable of making sense of your story? How to encode a combination of abstract contexts and emotions for an AI agent? Would that enable a more holistic contextual evaluation and better-informed decision-making process? We hypothesize that, by investigating abstraction and emotions traveling different spaces, we will gather insights into what abstract and emotion elements are key for a holistic and consistent understanding of emotion-rich contexts. Hence, we investigate the question: *How do abstraction and emotions travel different spaces?* We created a five-phase project (see fig. 1 for an overview) that shares similarities with the telephone game, as a message travels different spaces. Phases are crafted to gather insights into hidden or implicit abstract and emotion elements that help humans interpret emotion-rich contexts holistically. Zheng et al. [4] summarize the work described in [5], helping to describe implicit knowledge: "When knowledge has been articulated, then it is explicit knowledge. Otherwise, another question is raised: Can it be articulated? If the answer is yes, then it is implicit knowledge. If the answer is no, then it is tacit knowledge". Our project Phases are:

• Phase 0, **Image Collection**. Manually collect images whose sense-making requires a holistic approach due to abstract and emotion-rich contexts. We identified the digital space as a good fit to our purposes due to the social nature of emotions and their central place in digital cultures. That led us to images that convey jokes or metaphors characteristic in memes [6]. "The socially mediated communication of emotion is intricately linked to the social textures of networking technologies" [7].

• Phase 1, **Description and Data**. Write raw and detailed image descriptions and categorize the images in a dataset. Feed the Phase 2 team with raw descriptions. Example of a raw description: *A* soaking wet cat sits inside a sink with open eyes that pop out. There is a leading text: "I leave the bathroom shaking cold, and the person asks:" follow-up text: "'Are you cold?' Nope, a ghost is entering me." Unlike detailed, raw descriptions leave details out, which we named "unsaid elements" – Phase 1 team feeds Phase 2 with raw descriptions only, no images, and neither team uses generative AI tools.

We hypothesize that a) we will uncover hidden relationships across images' concrete and abstract elements as we categorize them, and b) our methods will provide insights into identifying diagnostic images [8] and providing emotion-aware assistive technologies.

• Phase 2, **3D Scenes and Decisions**. Without access to the images, interpret and encode the raw image descriptions into a 3D scene using a tool, such as Blender (see [9] for a review on Blender's versions and interfaces), and document the decision-making process. Unsaid elements can either be on the a) concrete side, e.g., it mentions a cat in a sink but no details about the fur's color or the sink's shape, size, and material/color; or b) more abstract, e.g., 3D modelers may reflect: "this seems to imply discomfort; is it supposed to be humorous?" Hence, 3D modelers have to fill in the gaps and make decisions to build a 3D scene – which we call "assumed elements". Therefore, *unsaid elements* from Phase 1 become *assumed elements* in Phase 2. In fig. 2, we illustrate the 3D modeler's decision-making process (illustration built using the Excalidraw tool [10]).

• Phase 3, **Checkpoint**. Compare: a) raw descriptions and images with 3D scenes and b) unsaid elements with assumed elements and documentation. Examine how/if those differ and what we learned about abstraction/emotions across spaces.

• Phase 4, **App.** Build an application to enable users to interact with our descriptions, dataset, and 3D scenes.

Project Phases and Teams. Each project phase has a dedicated team. A team is composed of people, and excluding Phase 4, they are forbidden to work on more than one project phase. Besides, teams have a strict non-sharing policy: everything a team produces is kept within the team only.

The paper is organized as follows: in Section I, we introduced our project idea; in Section II, we briefly provide background references as we discuss our project. Finally, we conclude in Section III.

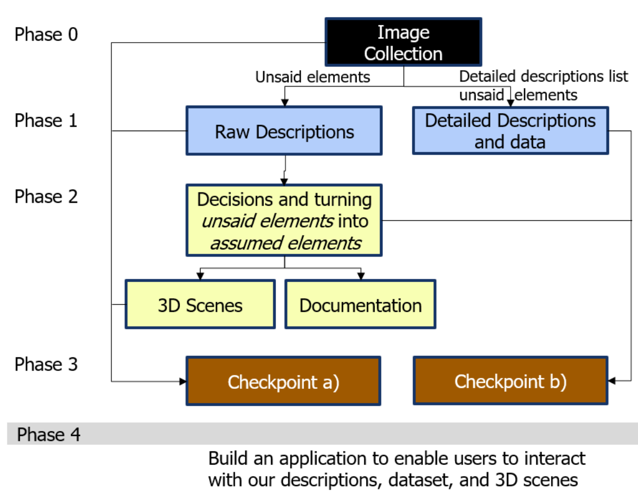

Fig. 1. An overview of the five project phases. For clarity, we omitted the arrows in Phase 4.

II. Background and Discussion

As we started collecting images for Phase 0, two topics became particularly relevant to our research: spatial communication and networked emotions. According to Tversky [11], by using position, form, and movement in space, gestures, and actions convey meanings. In that sense, differently from solely symbolic words, visual communication can directly convey content and structure (both literally and metaphorically). Although it may lack the rigorous definitions words can offer, visual communication delivers flexibility and suggestions for meanings. Such flexibility, in turn, requires context and experience to interpret conveyed meanings [11]. Networked emotions take into account the social nature of emotions and the messy layers of emotion and emotion regulation. It refers to the view of "emotions as multi-layered processes in which intraindividual processes are tightly coupled and often cannot be separated from interindividual processes" [12]. There are many instances where "regulation and elicitation can best be described by nested layers of feedback loops (...) Dealing with nested layers is messy because all layers can potentially influence emotional components" [12]. Finally, it "involves the mobilization of affect in online emotional cultures as a transmittable, spreadable, and self-contained resource, bringing out formerly privately shared emotions into online spaces and collective experience" [7].

As mentioned earlier, our project holds similarities with the telephone game. However, unlike the game, we are setting metrics to ensure objectivity without cutting off open-endedness. E.g., keeping consistent terminologies and processes across phases; building gateways for checking what unsaid elements tend to be correctly assumed by the Phase 2 team and which are not (by "correct," we mean objectively matchable with the original image); besides described concrete elements are easy to check across spaces. In short, the Phase 1 team examines Phase 0 images and produces raw textual image descriptions for the Phase 2 team, responsible for turning descriptions into 3D scenes. A comparison between *unsaid/assumed* elements and teams' documentation will help us investigate how abstraction and emotions travel through spaces. Interestingly, as Phases 0-2 start to shape, we notice a shift in our research questions from AI towards sense-making and networked emotions, and our applications moving to assistive technologies. We hypothesize that the mapping (unsaid/assumed elements) will help to better understand people's emotional mental models, and inform the development of emotion-aware AI systems and assistive technologies.

Our team members across phases engage in sense-making tasks which are, using Pirolli's and Card's [13] words: "information gathering, re-representation of the information in a schema that aids analysis, the development of insight through the manipulation of this representation, and the creation of some knowledge product or direct action based on the insight. In a formula Information → Schema → Insight → Product" [13]; and the re-representation may be in the team's mental models, written or drawn, or digitally represented.

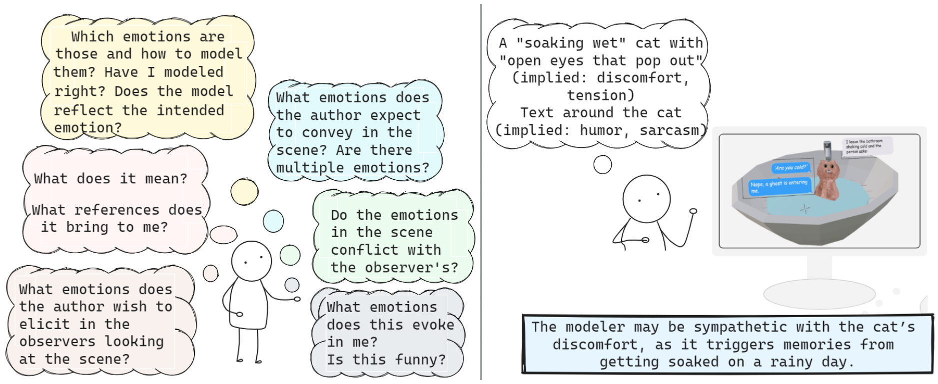

Fig. 2. Left. Challenges 3D modelers may face while dealing with multiple emotional-mental models. Although those are situated in the modeler's thoughts, they represent distinct players: modeler, 3D model, observer (or audience), description, and image's author). The figure illustrates 'layers' or dimensions of emotional processing involved in the decision-making and 3D modeling processes. Right. A modeler making decisions and designing a 3D scene.

Llorens-Gámez et al. [14] show that components, such as form and geometry, space distribution and context, color and texture, among others, influence memory and/or attention, and can be assessed objectively. It will be interesting to investigate, in Phase 2, to what extent familiar shapes or contexts populate a 3D modeler's assumed elements. If a modeler is used to seeing wood-made and square-like sinks, are those going to occupy matching assumed elements? (Of course, there are other players, such as how easy it is to design that shape and texture.) Or if an emotional context is related to disgust in the modeler's culture but anger in the original image's culture, will the 3D scene still be consistent with the original image? Images that are meant to be humorous to some may not be to others because humor shifts in different cultural contexts (see [15] for a view on how cultures create emotions). Modelers engage with networked emotions and emotional-mental models as they switch between and across mental models to guide the sense-making of a new description and decision-making that leads to creating a 3D scene. To conclude, this project raises many insightful questions for further investigation. For example, how to treat images that call for a "presupposed participant", images that expand their scope as they incorporate us, outside observers, as if we were part of the image/meaning? Back to AI systems, how to help an AI system to "see itself" as part of a context before producing a context's holistic understanding?

III. Conclusion

Seeking to understand what a holistic understanding of abstract and emotion-rich contexts could look like for an AI system, we created a five-phase project to investigate *how abstraction and emotions travel different spaces*. Given the challenges inherent to the investigation of emotions, we are identifying metrics to ensure objectivity and map how a message travels through spaces. As we do so, questions that are challenging but worth investigating are emerging, and we hypothesize they will bring insights into research in emotions and how to build emotion-aware AI systems and assistive technologies. For example, how informative would it be if an emotion-driven AI system outputted its decision log on emotions in a narrative-like sequence of pictures (or 3D scenes) and text?

Acknowledgments

We would like to thank Grinnell College's Harris Faculty Fellowship and the MAP program.

References

[1] K. L. van den Broek, J. Luomba, J. van den Broek, and H. Fischer, "Evaluating the application of the mental model mapping tool (m-tool)," *Frontiers in Psychology*, vol. 12, p. 761882, 2021.

[2] B. Tversky, "Cognitive maps, cognitive collages, and spatial mental models," in *European conference on spatial information theory*. Springer, 1993, pp. 14–24.

[3] B. Stangl, "Intrinsic mental models," in *Encyclopedia of the Sciences of Learning*. Springer, 2012.

[4] J. Zheng, M. Zhou, J. Mo, and A. Tharumarajah, "Background and foreground knowledge in knowledge management," in *International Working Conference on the Design of Information Infrastructure Systems for Manufacturing*. Springer, 2000, pp. 332–339.

[5] F. Nickols, "The tacit and explicit nature of knowledge: The knowledge in knowledge management," in *The knowledge management yearbook 2000-2001*. Routledge, 2013, pp. 12–21.

[6] R. Dawkins, *The selfish gene*. Oxford university press, 1976.

[7] K. Giaxoglou, K. Döveling, and S. Pitsillides, "Networked emotions: Interdisciplinary perspectives on sharing loss online," pp. 1–10, 2017.

[8] Y. Bai and W. Bainbridge, "Diagnostic images for alzheimer's disease show distinctions in biomarker status and scene-related functional activity between patients and healthy controls," *Journal of Vision*, vol. 23, no. 9, pp. 5600–5600, 2023.

[9] L. Soni, A. Kaur, and A. Sharma, "A review on different versions and interfaces of blender software," in *2023 7th I. Conf. on Trends in Electronics and Informatics (ICOEI)*. IEEE, 2023, pp. 882–887.

[10] E. Team, "Excalidraw," 2024. [Online]. Available: https://excalidraw.com/

[11] B. Tversky, "Visualizing thought," in *Handbook of human centric visualization*. Springer, 2014, pp. 3–40.

[12] A. Kappas, "Social regulation of emotion: messy layers," *Frontiers in psychology*, vol. 4, p. 51, 2013.

[13] P. Pirolli and S. Card, "The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis," in *Proceedings of international conference on intelligence analysis*, vol. 5. McLean, VA, USA, 2005, pp. 2–4.

[14] M. Llorens-Gámez, J. L. Higuera-Trujillo, C. S. Omarrementeria, and C. Llinares, "The impact of the design of learning spaces on attention and memory from a neuroarchitectural approach: A systematic review," *Frontiers of Architectural Research*, vol. 11, no. 3, pp. 542–560, 2022.

[15] B. Mesquita, *Between us: How cultures create emotions*. WW Norton & Company, 2022.