

Cheap and Easy Open-Ended Text Input for Interactive Emergent Narrative

Max Kreminski
mkreminski@scu.edu
Santa Clara University
Santa Clara, USA

i want something fizzy

The singer: Go to stageArea

Min is at the bar. June is at the bar. Isaac is at the bar. The singer is near the stage. June has a beer.

Isaac: Fulfill june's order

Min is at the bar. June is at the bar. Isaac is at the bar. The singer is near the entrance. June has a beer.

June: Order beer

Min is at the bar. June is at the bar. Isaac is at the bar. The singer is near the entrance. June is waiting for Isaac to serve them a beer.

Figure 1: A screenshot of our demo of PWIM. The text area in the top left allows the player to describe what they want to do; possible actions are presented below as clickable buttons, colored and sized to highlight actions that are similar to the input.

ABSTRACT

We present a demonstration of Play What I Mean (PWIM): a novel, AI-supported interaction technique for interactive emergent narrative (IEN) games and play experiences. By assisting players in translating high-level gameplay intents (expressed as short, unstructured text strings) into concrete game actions, PWIM aims to support open-ended player input while mitigating the overwhelm that players sometimes feel when confronting the large action spaces that characterize IEN gameplay. In matching player intents to game actions, PWIM makes use of an off-the-shelf sentence embedding model that is lightweight enough to run locally on a player’s device, and wraps this model in a simple user interface that allows the player to work around occasional classification errors.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; • **Applied computing** → **Computer games**; • **Computing methodologies** → *Natural language processing*.

KEYWORDS

interactive emergent narrative, interactive drama, sentence embeddings, interaction techniques

1 PLAY WHAT I MEAN

Interactive emergent narrative (IEN) games and play experiences, such as *Dwarf Fortress* and *The Sims*, allow players to perform actions in a wide-open sandbox that then responds narratively to whatever the player has done [7, 9]. The wide-open nature of player input in IEN supports many of the form’s central aesthetics [15, Sec. 3.2], but also tends to lead to player *overwhelm* [8] at the sheer breadth of actions that can be performed. Avoiding overwhelm while retaining open-endedness of input thus stands out as a central design challenge for IEN games.

Overwhelm at the breadth of the action space in IEN is partly downstream of user interface design: a graphical UI that makes all available game actions visible to the player, when backed up by a very wide action space, tends to appear visually overwhelming. For instance, the Versu UI [3]—which displays all available actions as distinct, clickable buttons—is sometimes described as confusing, due to the very large number of actions that must be explicitly considered at each choice point.

Alternatively, interactive narrative games with wide action spaces have often made use of a text parser to interpret commands entered by the player as short text strings, allowing players to access the full breadth of a wide-open action space without navigating an

overwhelming GUI. But parsers have problems of their own: verbs in parser interactive fiction are notoriously hard to discover [6], the syntax to which parser-interpreted instructions must conform is notoriously finicky [5], and the implementation of a truly forgiving parser for open-ended utterances (such as that in *Façade* [10], which aims to respond to arbitrary player-entered dialogue) represents a major investment of development time and effort—so major that very few games have ever shipped with such a parser.

To address these difficulties, we demonstrate a new interaction technique for IEN that aims to provide some of the advantages of open-ended text input without necessitating the development or reuse of a sophisticated parser. This technique—Play What I Mean (PWIM)—builds on the basic *text-to-dialogue* approach proposed by Chen et al. [1] (i.e., using a sentence embedding model to map player-entered intent strings to game-provided lines of player character dialogue) by combining it with a more sophisticated, Versu-inspired system for state tracking and conditional action provision [2], as well as a user interface that enables players to explicitly choose between several game actions that represent *likely* matches for their intent. Our demo of PWIM can be played in a web browser¹ and is implemented completely in clientside JavaScript using a generic, off-the-shelf sentence embedding model [12], demonstrating the ease of implementation of this interaction technique for arbitrary IEN games.

2 IMPLEMENTATION

The core of the PWIM interaction loop is as follows:

- (1) The player types a natural language phrase or sentence describing what they want to do.
- (2) Currently available game actions are sorted according to semantic similarity to this sentence.
- (3) Sorted actions are displayed in a way that visually prioritizes closer matches, but allows the player to choose what action is ultimately performed.

In our demonstration, possible game actions are provided by a Praxish domain that mirrors the functionality of Versu. Praxish actions are defined as small blobs of structured JSON, making them relatively straightforward to author; see Fig. 3 in the original Praxish paper [2] for an example.

Semantic similarity between available game actions and a player-entered intent phrase is determined via the `all-mpnet-base-v2` model provided with the open-source `SentenceTransformers` library [12]. Every available action is summarized as a short imperative phrase, and embeddings of these action summary phrases are compared to an embedding of the player intent phrase via cosine similarity. Closer matches are displayed with a proportionally darker background color in the PWIM UI (Fig. 1), and the top K matches (in our demo, $K = 3$) are displayed at an enlarged size to further increase their visual prevalence.

One way to explore the potential of this approach is to travel to the bar in our example game, then enter a wide variety of intent phrases. In our testing, we found that phrases as diverse as “get hammered”, “gimme something autumnal”, “play music”, “gotta stay hydrated”, and “say hisaac” all resulted in the most closely aligned possible action being surfaced as the top recommended action in

the PWIM UI. Meanwhile, only a few testing phrases (e.g., “sober up”) did not result in appropriate top recommendations. Occasional misclassification of intent phrases is why we choose to display all possible actions in the UI (giving players a chance to manually correct misclassifications) instead of automatically performing the action that represents the closest match; misclassifications can also likely be mitigated by finetuning the sentence embedding model against a game-specific set of player intent and action summary phrases, although this would increase deployment complexity.

3 RELATED WORK

Chen et al. [1] have previously employed a similar approach to ours, using sentence embeddings to map player utterances to pre-recorded lines of dialogue. However, this earlier work does not tie dialogue selection into a deeper emergent narrative simulation with complicated state tracking and many conditionally available actions. Additionally, player intent phrases are used in this work to automatically play the most closely matching line of dialogue; this simpler approach is more susceptible to derailment by misclassification of intent phrases, and may limit the discoverability of game actions in comparison to PWIM.

Some co-creative narrative play experiences have also made use of purely language modeling approaches—including data-driven case-based reasoning systems [13, 17], text-generating recurrent neural networks [14], and large language models [4, 16]—to respond to arbitrary text input by players. However, these systems generally do not attempt to ground out player input in a hand-crafted ontology of game actions, often limiting the coherence of the stories that these systems can produce—or, alternatively, requiring the intervention of an external game master to weave the results of command invocations into a coherent story [11].

ACKNOWLEDGMENTS

Thanks to Barrett R. Anderson for initially exploring the possibility of moving the semantic similarity component of the PWIM demo into clientside JavaScript.

REFERENCES

- [1] Sherol Chen, Ben Pietrzak, Kory Mathewson, Anna Kipnis, and Erin Hoffman-John. 2020. Text to Dialog: Using Semantic Similarity to Extend Narrative Immersion in Virtual Worlds. In *NeurIPS Workshop on Machine Learning for Creativity and Design 4.0*.
- [2] James Dameris, Rosaura Hernandez Roman, and Max Kreminski. 2023. Praxish: a rational reconstruction of a logic-based DSL for modeling social practices. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 407–415.
- [3] Richard Evans and Emily Short. 2014. Versu—a simulationist storytelling system. *IEEE Transactions on Computational Intelligence and AI in Games* 6, 2 (2014), 113–130.
- [4] Minh Hua and Rita Raley. 2020. Playing With Unicorns: AI Dungeon and Citizen NLP. *DHQ: Digital Humanities Quarterly* 14, 4 (2020).
- [5] IFWiki Contributors. 2006. Infocom-type parser. https://www.ifwiki.org/Infocom-type_parser. Accessed: 2023-12-29.
- [6] IFWiki Contributors. 2017. Guess-the-verb. <https://www.ifwiki.org/Guess-the-verb>. Accessed: 2023-12-29.
- [7] Max Kreminski and Michael Mateas. 2021. A coauthorship-centric history of interactive emergent narrative. In *Interactive Storytelling: 14th International Conference on Interactive Digital Storytelling, ICIDS 2021, Tallinn, Estonia, December 7–10, 2021, Proceedings 14*. Springer, 222–235.
- [8] Max Kreminski, Noah Wardrip-Fruin, and Michael Mateas. 2023. Authoring for story sifters. In *The Authoring Problem: Challenges in Supporting Authoring for Interactive Digital Narratives*. Springer, 207–220.

¹<https://mkremins.github.io/pwim>

- [9] Sandy Louchart, John Truesdale, Neil Suttie, and Ruth Aylett. 2015. Emergent narrative: past, present and future of an interactive storytelling approach. In *Interactive Digital Narrative*. Routledge, 185–199.
- [10] Michael Mateas and Andrew Stern. 2003. Façade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference*.
- [11] Stefan Papazov, Wesley Gill, Marta Garcia Ferreiro, Andrew Zhu, Lara J Martin, and Chris Callison-Burch. 2022. Using language models to convert between natural language and game commands. In *The Third Wordplay: When Language Meets Games Workshop*.
- [12] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics. <https://arxiv.org/abs/1908.10084>
- [13] Melissa Roemmele and Andrew S Gordon. 2015. Creative Help: A story writing assistant. In *Interactive Storytelling: 8th International Conference on Interactive Digital Storytelling, ICIDS 2015, Copenhagen, Denmark, November 30-December 4, 2015, Proceedings 8*. Springer, 81–92.
- [14] Melissa Roemmele and Andrew S Gordon. 2018. Automated assistance for creative writing with an RNN language model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*.
- [15] James Ryan. 2018. *Curating Simulated Storyworlds*. Ph. D. Dissertation. University of California, Santa Cruz.
- [16] Yuqian Sun, Zhouyi Li, Ke Fang, Chang Hee Lee, and Ali Asadipour. 2023. Language as reality: a co-creative storytelling game experience in 1001 Nights using generative AI. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 19. 425–434.
- [17] Reid Swanson and Andrew S Gordon. 2012. Say Anything: Using textual case-based reasoning to enable open-domain interactive storytelling. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 3 (2012).