Scaling Analysis of Creative Activity Traces via Fuzzy Linkography

Amy Smith
Queen Mary University of London
London, United Kingdom
amy.smith@qmul.ac.uk

Barrett R. Anderson
Independent Researcher
San Mateo, California, USA
barrettrees@gmail.com

Jasmine Tan Otto
Independent Researcher
San Mateo, California, USA
ottojasmine@gmail.com

Isaac Karth
Independent Researcher
Berkeley, California, USA
isaac@isaackarth.com

Yuqian Sun
Midjourney
London, United Kingdom
ysun@midjourney.com

John Joon Young Chung
Midjourney
San Francisco, California, USA
jchung@midjourney.com

Melissa Roemmele
Midjourney
San Francisco, California, USA
mroemmele@midjourney.com

Max Kreminski
Midjourney
Berkeley, California, USA
mkreminski@midjourney.com

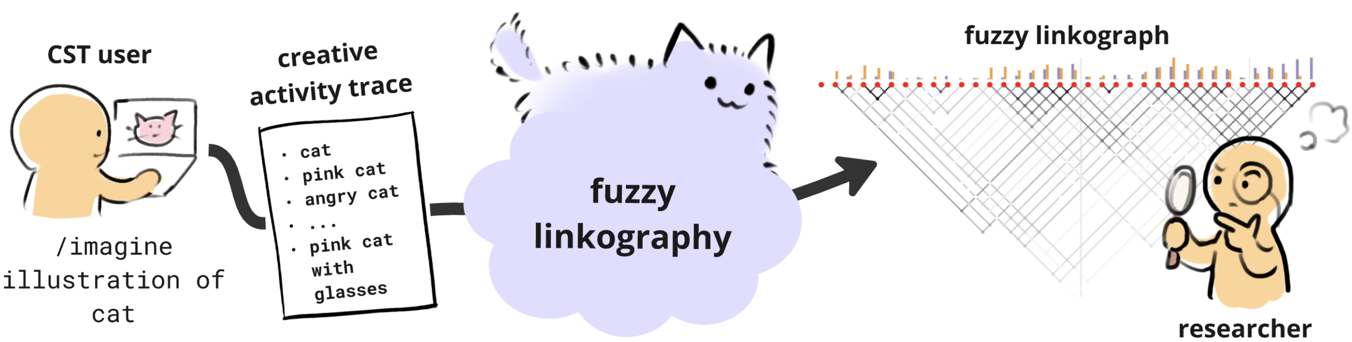

Figure 1: Fuzzy linkography allows for the rapid translation of user activity logs from digital creativity support tools (and other traces of creative activity) into rough graphical summaries, suitable for visual and quantitative inspection by researchers.

Abstract

Creativity researchers sometimes employ *linkography*—a family of techniques involving the visualization and analysis of links between *design moves*—to make sense of people's behavior in creative contexts, but traditional linkography (which involves manual annotation of both moves and links) is so time-consuming that it is mostly applied at very small scales. Meanwhile, digital creativity support tools (e.g., text-to-image prompting tools) automatically capture vast numbers of user interaction traces, but these traces are not yet well understood. We introduce a means of quickly and automatically producing *fuzzy linkographs* of text-to-image prompting traces and apply this technique to a large corpus of traces collected from the live deployment of a commercial text-to-image tool. This allows us to uncover recurring linkographic structures in text-to-image prompting interactions; cluster traces to classify user behaviors into several distinct archetypes; and quickly sift through thousands of traces to surface structurally interesting episodes of prompting.

CCS Concepts

• **Human-centered computing** → **Visualization techniques**; *Empirical studies in HCI*; • **Applied computing** → *Arts and humanities*; • **Computing methodologies** → Natural language processing.

Keywords

creativity support tools, interaction dynamics, protocol studies of design, evaluation methods, visualization, sentence embedding

ACM Reference Format:
Amy Smith, Barrett R. Anderson, Jasmine Tan Otto, Isaac Karth, Yuqian Sun, John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Scaling Analysis of Creative Activity Traces via Fuzzy Linkography. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3706599.3720184

CHI EA '25, April 26-May 1, 2025, Yokohama, Japan

ACM ISBN 979-8-4007-1395-8/2025/04
https://doi.org/10.1145/3706599.3720184

1 Introduction

Digital creativity support tools (CSTs) [11, 17, 57]—software systems intended to support users' creative activity—have become increasingly prevalent in the last few decades, both in professional and casual contexts [13]. Recent years have seen especially strong growth in the adoption of AI-based CSTs—in creative domains as wide-ranging as writing [36], visual art [7, 45], game design [23], fashion design [28], worldbuilding [12], and music [40].

Widespread use of digital CSTs permits the automatic capture of traces of human creative activity at a previously unprecedented scale, which could be of assistance both in developing theories of creativity and in evaluating CSTs. Evaluating CSTs is difficult [54] and has been known to be difficult since the first days of CST research [26]; many evaluation approaches focus on the assessment of CST users' subjective experiences of the creative process [8, 35], while others focus more on the products of CST use [2, 41]. However, recording of user interaction traces also permits the evaluation of (co-)creative *interaction dynamics*, including through the visualization and analysis of user trajectories through design space [1, 14, 16, 32]. These evaluations are process-focused and attempt to characterize how creative activity evolves over time, allowing them to directly inform process models of creativity [21, 42]—but their applicability has so far been limited by the need for domain-specific quantitative metrics to automatically characterize user-created artifacts [1, 32] or manual qualitative coding of video recordings of interaction [14, 16] as prerequisites for use.

Linkographic analyses [18, 19] are similarly visual, temporal, and process-focused, and their generality has enabled application to a wide range of creative domains [6, 10, 15, 24]. Linkography hinges on annotation and visualization of the *design moves* that make up an episode of creative activity and the *links* between related moves; various statistics can also be computed on the resulting graphs [19, 30]. Although linkography is sometimes employed to investigate user interactions with CSTs [55], including AI-based CSTs [37], it remains a "costly research method, both in terms of time and resources" [51]; as a result, linkography is still rarely used to analyze large numbers of creative activity traces captured by digital CSTs, and some researchers even avoid linkography at smaller scales due to its "logistical and labour overheads" [3].

To address this limitation, we introduce **fuzzy linkography**: a technique for automatically producing (imperfect) linkographs from sequences of recorded design moves, using a computational model of semantic similarity as a stand-in for the human annotator of links between design moves. We apply our approach to text-to-image prompting journeys [44] and highlight recurring linkographic motifs in these journeys at scale, suggesting that our approach can be used to discover interesting structure in a variety of human-AI creative interactions in the future.

2 Linkography: A Brief Primer

Traditional linkography has been explicated in a number of prior publications (e.g., [19, 25, 30, 37, 51]), so we do not attempt to describe it fully here. Instead, we first present a brief overview of key terms and concepts from the linkography literature. Section 3 then updates these concepts for use with automatically constructed linkographs of continuously weighted rather than binary links.

A linkograph represents a single *design episode* and consists of two primary components: a sequence of *design moves* and a set of pairwise *links* between these design moves. **Design moves** represent concrete changes made to the *design situation*, and are conventionally plotted left to right, with each move represented by a single dot and uniform spacing between moves (regardless of how much time elapsed between each pair of moves). **Links** represent human-annotated connections between related design moves: if a later design move can be said to *build on* an earlier design move, a link is drawn connecting these two moves. Links are conventionally drawn below the move sequence.

Several statistics can be computed on linkographs. Individual moves have **backlink** and **forelink counts** that indicate how many other moves they built on and influenced respectively. Moves with an especially high link count are called **critical moves** (CMs); forelink CMs are sometimes seen as *divergent* and backlink CMs are sometimes seen as *convergent.* A **link density index** (LDI) for the linkograph as a whole can be calculated by dividing the total number of links by the total number of moves; this value indicates the overall interconnectedness of the moves in the graph. To quantify the unpredictability of links in a graph (as a rough proxy for the dynamism of the whole design episode), several measures of **link entropy** are also sometimes employed; these calculations are complicated to describe, and we do not make much use of them here. See Appendix A for details.

Visually, linkographs are often analyzed in terms of the structures that appear within them. Goldschmidt [19] describes three key pattern types that may be seen in linkographs (Figure 2): **chunks**, or sequences of interrelated moves that begin with a clear "inciting" move and are mostly linked back to that incident; **webs**, or more tightly interrelated clusters of moves in which each move is linked to almost every other; and **sawtooths**, or sequences of moves in which each move is related only to its immediate predecessor and successor (potentially suggesting the development of a single idea that is largely not tied into the rest of the design situation). Linkless moves are called **orphans** and indicate ignored digressions.

3 Fuzzy Linkography

Fuzzy linkography is much like traditional linkography, with two key differences. First, links between moves are *automatically inferred* by a computational process rather than manually annotated by a human coder. Second, links are represented as numbers ranging from zero to one (indicating the **strength** of semantic association between a pair of moves) rather than binary on/off values. Continuous link strength values, and their imperfect correlations with human assessments of move relatedness, are what give fuzzy linkographs their fuzziness; rather than forcing the machine annotator to make an authoritative-seeming binary choice about whether each pair of moves is or is not related, we prefer to pass information about move associations that the machine finds ambiguous or uncertain along to the human user of linkography. For a detailed walkthrough of our visual notation for fuzzy linkographs, see Figure 9; for an extended discussion of limitations, see Appendix B.

In our implementation of fuzzy linkography, links between moves are established via an *embedding model* [53] that translates textual descriptions of design moves into vectors. For each pair of moves,

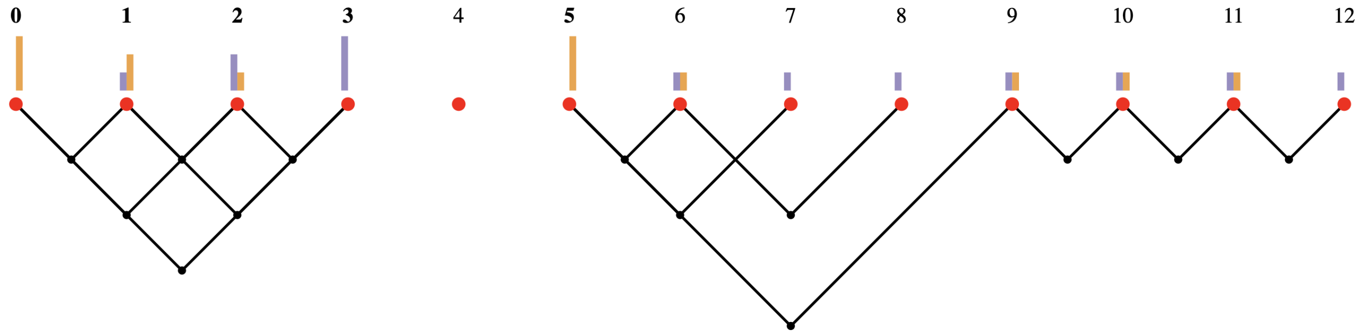

Figure 2: Standard linkographic patterns. Moves 0–3 form a *web*; move 4 is an *orphan*; moves 5–9 form a looser *chunk*; and moves 9–12 form a *sawtooth.* Above each move, purple and orange bars indicate the move's *backlink* and *forelink count* respectively; moves with an especially high link count in either direction may be deemed *critical moves,* e.g., the *forelink critical move* 5.

we use cosine similarity between the embedding vectors representing each move to determine the strength of the link between these moves. A numeric similarity threshold t is used to discard weak associations between moves, so that not all moves are judged as being linked to some extent; raw cosine similarity values exceeding this threshold are then linearly rescaled from the range $[t, 1]$ to the range $[0, 1]$ to establish link strengths.

When rendering fuzzy linkographs, we use link color—ranging from white, for very weak links, to black, for very strong links—to indicate the strength of each link. "Fuzzy" lighter-colored links thereby serve to communicate model uncertainty about move association; visual "fuzziness" has been found to be an intuitive way to convey uncertainty [50], which can help to improve the transparency of analyses that rely on machine learning models [4, 5].

Because fuzzy links are represented as numeric strength values rather than binary on/off states, the quantitative measures calculated in traditional linkography must be updated to work with fuzzy linkographs. Forelink and backlink weight values on a fuzzy linkograph can be calculated simply by summing the strengths of a move's forelinks and backlinks respectively. Link density index values can similarly be computed by summing the strengths of all links in the graph and dividing this by the total number of moves. Entropy values can be updated similarly; see Appendix A.

To support widespread application of fuzzy linkography, we open-source the code we use to construct, visualize, and quantitatively analyze linkographs.[1] All analyses that we report on in this paper assess design move similarity via the `all-MiniLM-L6-v2` sentence embedding model [53]—an open-source, open-data and open-weights model that has previously been validated against a human baseline for assessment of semantic similarity in the context of open-ended ideation [2]. For all analyses reported in this paper, we use a fixed similarity threshold $t = 0.35$ as the minimum similarity score for which we infer a link; this value seems to work well with our chosen embedding model, although in the future, it may be worth investigating whether the value of t can somehow be chosen in a more principled way.

[1] https://github.com/mkremins/fuzzy-linkography

4 Analyzing Image Prompting Journeys

We apply fuzzy linkography to the analysis of user image prompting journeys in a graphical creativity support tool built around a popular commercial text-to-image diffusion model. Image prompting practices have previously been analyzed computationally via topic modeling [46, 56], but we are unaware of any prior large-scale examination of how individual users' prompts develop over time. Our dataset consists of 6,424 interaction traces documenting every image prompt submitted by every user of the tool during the period from Dec 11–31, 2024. To exclude data from users who did not make much use of the tool, we filtered these interaction traces to 1,879 "substantial" traces that contain at least seven image prompts each. The longest trace in our filtered dataset is 536 moves long, and the filtered traces overall have a median length of 14.

We consider each interaction trace as a separate design episode and construct a fuzzy linkograph of the episode, treating individual textual prompts submitted by the user as design moves. Our (unoptimized) Python implementation of link inference, running in Python 3.9.6 on an Apple MacBook Air (M2, 2022), takes 752.944 seconds to compute link strengths for the filtered set of traces—roughly 0.4 seconds per trace. We then compute several linkographic statistics on each trace, including link density index; move-level forelink and backlink weights (which can be used to help identify critical moves); graph-level forelink, backlink, and horizonlink entropy values; and an overall link entropy value summing up the other entropies.

4.1 Recurring Linkographic Motifs

By visually inspecting linkographs, noting recurring patterns, and investigating the specific prompts that are involved in these patterns, we can build a taxonomy of structural *motifs* that frequently appear in image prompting traces.

4.1.1 Refinement Webs. The most common motif in our image prompting linkographs is a "web" of tightly interconnected moves, signifying a moment of prompt *refinement*: the user gradually testing smaller and smaller variations on a prompt with relatively fixed subject matter as they narrow in on the images they want. Links within this web can often be seen to get gradually stronger from the beginning to the end of the sequence, as the submitted prompts

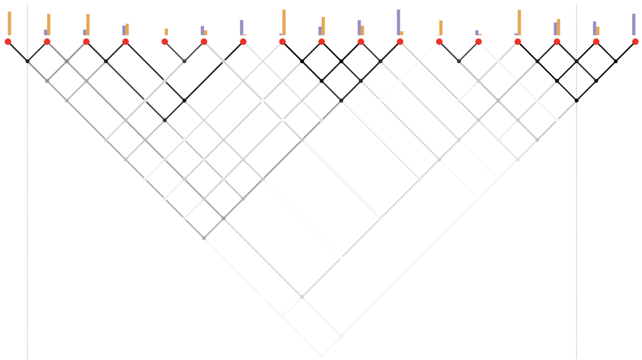

Figure 3: This image prompting linkograph contains three refinement webs: one at the beginning, one in the middle, and one at the end. The first web consists of mixed medium-strength and high-strength links, indicating a greater degree of prompt variation at an earlier level of conceptualization; the other two webs mostly consist of very small permutations of a relatively fixed prompt.

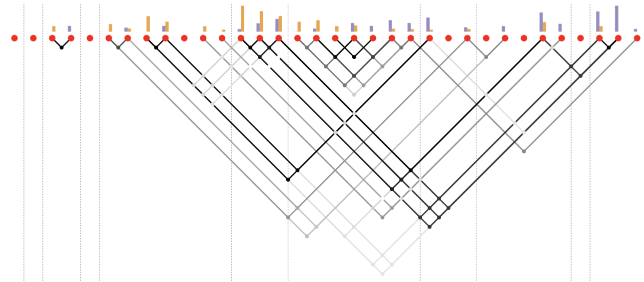

Figure 4: This image prompting linkograph demonstrates a curiosity zigzag, with relatively strong long-range links indicating returns to a central theme and disconnected chunks between these moves indicating periods of exploration.

become steadily more similar to one another—eventually often culminating in several retries of the exact same prompt. In Figure 3, for example, three refinement webs correspond to a user's attempts to visualize an alien civilization; a particular location on that civilization's home planet; and members of a specific faction within that civilization, respectively.

Contrasting somewhat with an earlier finding that a prompting-based interface usually resulted in relatively small webs compared to a sketching-based interface [37], we found that large webs of tightly interconnected moves are fairly common in our text-to-image linkographs. We suspect this may be because the interface we studied generates multiple different images in response to a single prompt submission, perhaps increasing the perceived value to users of retrying identical or near-identical prompts to stochastically explore a space of possible results.

4.1.2 Curiosity Zigzags. Another common motif in our image prompting linkographs is a large-scale "zigzag" structure temporally interspersed with smaller and largely unrelated chunks or webs, often corresponding to user alternation between a "central theme" that they keep returning to and a set of further-flung explorations of different ideas or themes. In Figure 4, for instance, a user periodically returns to a single central subject—a cyclopean "black spirit" character—between largely unrelated and wide-ranging explorations of other themes.

This mirrors a pattern observed in an earlier visualization-based study of user trajectories through a casual CST's design space [32], in which some users visibly alternated between outputs sampled from a relatively narrow "home base" of outputs and further-flung explorations that push the boundaries of the CST's expressive range [59]. Relative lack of integration of further-flung explorations may be taken as indicative of boundary-pushing, probing the CST's capabilities, or exploration for its own sake (perhaps indicative of the "curious users of casual creators" phenomenon [47] detected in even older work); in some cases we believe this pattern may also indicate user non-development of a coherent goal.

4.1.3 Zigzags Toward Convergence. One recurring motif combination, which we term the "converging zigzag", involves a curiosity zigzag with gradually growing integration between the "main theme" and divergent exploration chunks or webs over time (Figure 5). This structure seems to be especially characterized by refinement webs of progressively strengthening links at the "tail end" of a move sequence, corresponding to integrations of previously unrelated themes.

Figure 5 shows two such webs. The first web corresponds to a gradual drawing-together of two recurring *subject* themes in this user's prompts (an insect and a "cute old man" character); the second corresponds to a drawing-together of this hybrid character subject with a visual *style* that the user had previously explored in other, unrelated images.

Interestingly, the first refinement web occurs just before the user stops engaging with the CST for a while. The second refinement web also occurs near two temporal "session breaks"; it crosses over the first break and terminates with the second, indicating that the user did not engage further with the CST following their completion of the second web. This seems to indicate that the user twice gradually formed a specific goal and progressed toward its realization, then stepped away from the CST once the goal was realized. More broadly, across our image prompting data, refinement webs toward the tail end of a user activity period generally seem to indicate the user forming and realizing a concrete goal, then stepping away with the "final" generated images.

4.1.4 Variations in Temporal Structure. We observed clear differences between users in terms of how their activity changed across breaks from prompting. The two most obvious temporal patterns we noticed in linkograph structure are temporally-divided *sessions* (Figure 6) and temporally non-divided *projects* (Figure 7). Sessions are chunks of related prompting activity that are cleanly bounded by the user taking a break from prompting on each side; projects are chunks of related prompting activity that span multiple such breaks. We were initially surprised to see such clear persistence of prompting subjects across temporal divides.

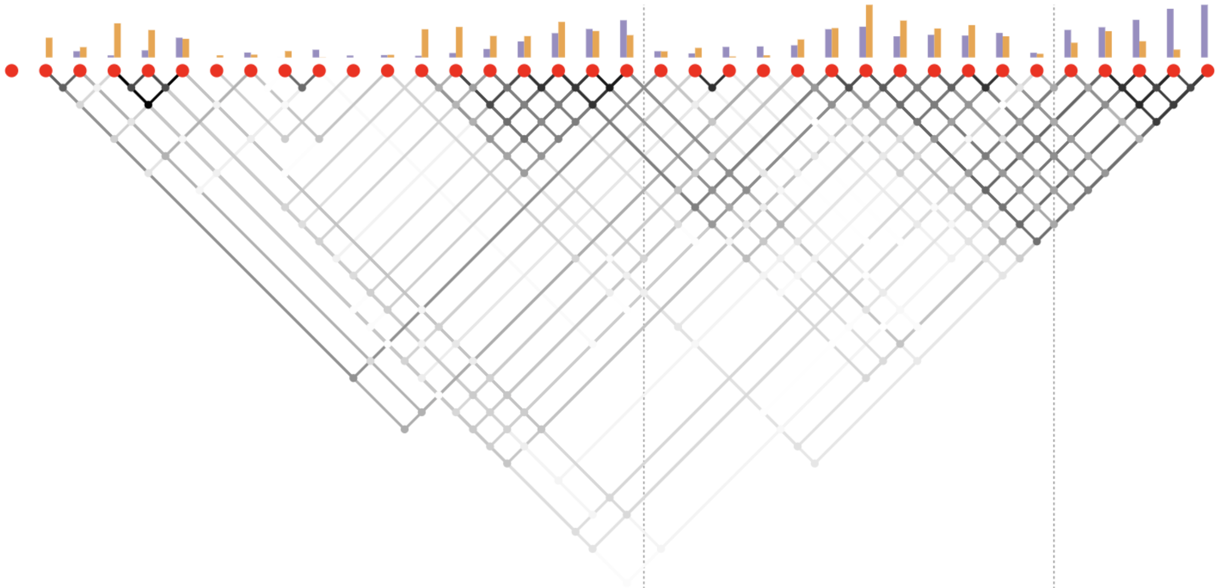

Figure 5: This image prompting linkograph's "converging zigzag" shape combines a *curiosity zigzag* with *refinement webs*: alternation between a central theme of interest and more divergent explorations persists throughout the episode, but previously distinct themes are also brought together into a single image concept on two distinct occasions, each time resulting in a distinctive refinement web.

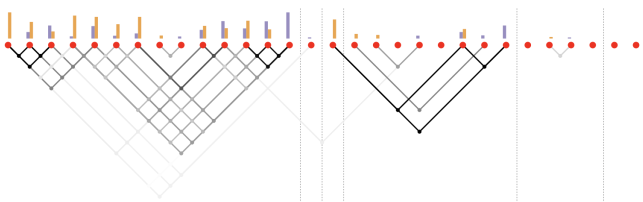

Figure 6: This image prompting linkograph shows clear semantic divisions between the user's focal themes during several significant temporal "sessions" of prompting. (Vertical dotted lines between moves indicate the user stepping away from prompting for at least 30 minutes.) The first and longest session ends with a clear refinement web.

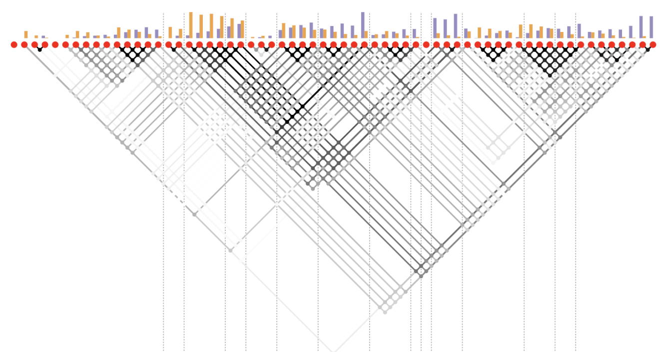

Figure 7: This image prompting linkograph, on the other hand, shows a user's focal themes persisting across multiple clear temporal divides in prompting activity. The user seems to be engaged in a long-running "project".

4.2 Trace Clustering

Computing linkographic statistics on each trace also allows us to *cluster* the traces, associating each user's activity with a distinguishable archetype. We first translate each trace into a *signature vector* of three statistics: move count, link density index (LDI), and overall link entropy. We then normalize these values to z-scores and discard as an outlier any trace whose signature vector includes a z-score greater than 3; this excludes 59 traces, most of which are unusually long (consisting of hundreds of moves) and thus probably in need of partitioning for further analysis. We cluster the remaining traces' signature vectors using k-means clustering—as employed for design style clustering [1]—with $k = 5$.

The clusters we encounter (Figure 8) can be described as follows:

(1) Medium-long episodes with multiple distinct refinement webs *(331 traces)*
(2) Short-medium episodes of mostly disconnected ideas *(729 traces)*
(3) Short-medium episodes containing a strong refinement web *(273 traces)*
(4) Long episodes of very densely interconnected moves *(174 traces)*
(5) Medium-long zigzaggy episodes with mostly short-range links but some longer-range links too *(313 traces)*

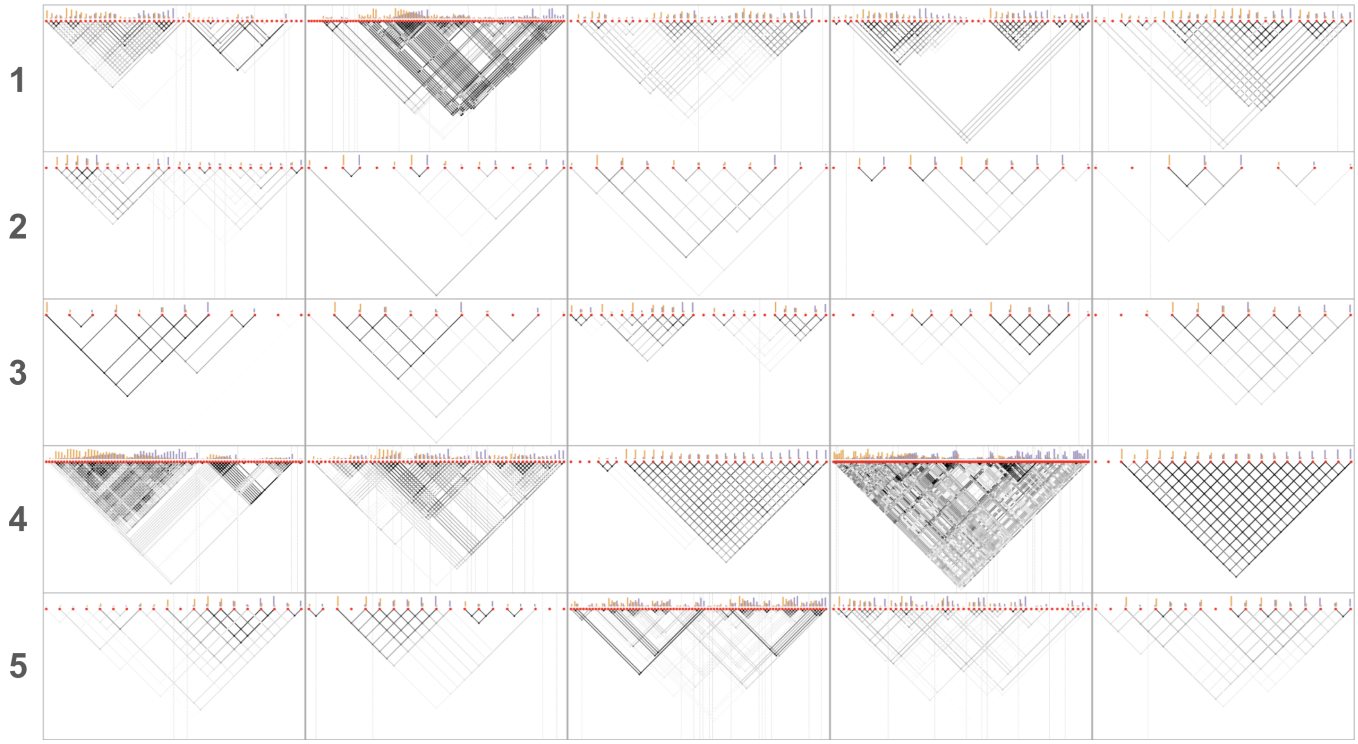

Figure 8: Five clusters of image prompting linkographs, five random examples per cluster.

These clusters, although relatively rough, demonstrate the potential of *designer modeling* [39] via automatically associating CST users with distinguishable user archetypes based on their linkographic activity patterns. For instance, members of cluster 1 often seem to repeatedly seek out an interesting region of image space, then gradually refine a single prompt until they converge on an image they like; members of cluster 2 mostly prompt for unrelated or largely unrelated concepts in batches of one or two prompts at a time; and members of cluster 4 often seem to have a single major *fascination* and mostly prompt for images of this fascination.

Tracking the relative frequency of traces in each category might help guide design iteration on a deployed CST: for instance, a high frequency of short and low-connectivity episodes may indicate that users need additional scaffolding for goal formation. Meanwhile, cluster-based classification of individual users might allow in-the-moment adaptation of a CST's user interface to user-archetypal needs: for instance, users in cluster 5 may find more value than others in features involving the revisitation and recombination of past prompting subjects or themes.

5 Discussion

Potential applications of fuzzy linkography go far beyond those investigated here. Large-scale application may be used to *verify linkographic hypotheses* about what kinds of linkographs indicate a "good" episode of design or ideation (e.g., [29]), or about how creativity functions as a process (e.g., [20]). Quantitative metrics on fuzzy linkographs could also be used for gauging creative momentum [31], to decide when would be a good time for the AI system to proactively intervene in a co-creative interaction; or to identify past moves that are as yet "undeveloped", to focus proactive generation on unexplored potentialities.

Realtime generation of linkographs opens the possibility of displaying linkographs during a design episode to the participants as a form of reflective visualization. Reflective visualization has proven helpful in co-creative contexts in the past [38] and some CSTs are explicitly designed to promote reflection [33, 43]; the display of fuzzy linkographs to users may be especially impactful in this context. However, the interpretability of fuzzy linkographs by untrained users remains unevaluated; more research may be necessary to validate broader deployment of fuzzy linkography among non-researchers.

Like homogenization analysis [2], fuzzy linkography can be applied to any artifacts that can be embedded in a semantic space—not just short texts but also images [52, 65], longer texts [60, 63], music [22, 27], user interfaces [64], 3D models [34, 61], game states [66], and likely others in the future. Alternative embedding models may even allow direct incorporation of design moves *by the machine* (e.g., generated images) into fuzzy linkographs. This could enable the analysis of connectivity patterns between human and machine moves in a human-AI co-creative context, allowing for identification of machine moves that strongly shape the human's ideation; cases of the human and machine "talking past one another" rather than integrating their ideas; and so on. We describe our first attempt at inter-actor connectivity pattern analysis in fuzzy linkographs featuring both human and machine moves in Smith et al. [58].

6 Conclusion

We have introduced a technique for the automatic construction of *fuzzy linkographs* from creative activity traces, including those collected naturally through the course of user interactions with digital creativity support tools. We have also demonstrated the application of this technique to text-to-image prompting interactions, discovering a variety of interesting linkographic patterns. Although the resulting linkographs are imperfect, they nevertheless function well as *graphical summaries* of design episodes, surfacing high-level structural patterns in creative activity traces at a glance and serving as jumping-off points for deeper analysis. In closing, we would especially like to stress the potential for linkographic *abundance* (Figure 10) to enable new applications of linkography: whereas linkographs have up until this point been relatively scarce, the availability of low-cost approaches to linkography may greatly expand the potential audience for linkographic techniques, for instance by allowing CST developers to rapidly create linkographs for very large numbers of user activity traces as a window into broad usage trends.

References

[1] Alberto Alvarez, Jose Font, and Julian Togelius. 2022. Toward Designer Modeling Through Design Style Clustering. *IEEE Transactions on Games* 14, 4 (2022), 676–686.

[2] Barrett R Anderson, Jash Hemant Shah, and Max Kreminski. 2024. Homogenization effects of large language models on human creative ideation. In *Proceedings of the 16th Conference on Creativity & Cognition*. 413–425.

[3] Jack Armitage, Thor Magnusson, and Andrew McPherson. 2023. Design Process in Visual Programming: Methods for Visual and Temporal Analysis. In *Proc. Sound and Music Computing Conference*.

[4] Jesse Josua Benjamin, Arne Berger, Nick Merrill, and James Pierce. 2021. Machine learning uncertainty as a design material: A post-phenomenological inquiry. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

[5] Umang Bhatt, Javier Antorán, Yunfeng Zhang, Q Vera Liao, Prasanna Sattigeri, Riccardo Fogliato, Gabrielle Melançon, Ranganath Krishnan, Jason Stanley, Omesh Tickoo, Lama Nachman, Rumi Chunara, Madhulika Srikumar, Adrian Weller, and Alice Xiang. 2021. Uncertainty as a form of transparency: Measuring, communicating, and using uncertainty. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. 401–413.

[6] Hui Cai, Ellen Yi-Luen Do, and Craig M Zimring. 2010. Extended linkography and distance graph in design evaluation: an empirical study of the dual effects of inspiration sources in creative design. *Design Studies* 31, 2 (2010), 146–168.

[7] Minsuk Chang, Stefania Druga, Alexander J Fiannaca, Pedro Vergani, Chinmay Kulkarni, Carrie J Cai, and Michael Terry. 2023. The prompt artists. In *Proceedings of the 15th Conference on Creativity and Cognition*. 75–87.

[8] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the Creativity Support Index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014).

[9] Shyan-Bin Chou. 2007. A method for evaluating creativity in Linkography. In *10th QMOD Conference of Quality Management and Organizational Development. Our Dreams of Excellence*, Vol. 26. 18–20.

[10] Shyan-Bin Chou, Huey-Wen Chou, and Yui-Liang Chen. 2013. Entropy of Linkography: Evaluating the creativity of short animation. *Creativity Research Journal* 25, 1 (2013), 33–37.

[11] John Joon Young Chung, Shiqing He, and Eytan Adar. 2021. The intersection of users, roles, interactions, and technologies in creativity support tools. In *Designing Interactive Systems Conference 2021*. 1817–1833.

[12] John Joon Young Chung and Max Kreminski. 2024. Patchview: LLM-Powered Worldbuilding with Generative Dust and Magnet Visualization. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*.

[13] Kate Compton and Michael Mateas. 2015. Casual Creators. In *International Conference on Computational Creativity*. 228–235.

[14] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative sense-making: Quantifying interaction dynamics in co-creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. 356–366.

[15] Stefano Delle Monache and Davide Rocchesso. 2016. Cooperative sound design: A protocol analysis. In *Proceedings of the Audio Mostly 2016*. 154–161.

[16] Manoj Deshpande, Jisu Park, Supratim Pait, and Brian Magerko. 2024. Perceptions of Interaction Dynamics in Co-Creative AI: A Comparative Study of Interaction Modalities in Drawcto. In *Proceedings of the 16th Conference on Creativity & Cognition*. 102–116.

[17] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*.

[18] Gabriela Goldschmidt. 1990. Linkography: assessing design productivity. In *Cyberbetics and System '90, Proceedings of the Tenth European Meeting on Cybernetics and Systems Research, Singapore*. World Scientific, 291–298.

[19] Gabriela Goldschmidt. 2014. *Linkography: Unfolding the Design Process*. MIT Press.

[20] Gabriela Goldschmidt. 2016. Linkographic evidence for concurrent divergent and convergent thinking in creative design. *Creativity Research Journal* 28, 2 (2016), 115–122.

[21] Adam E Green, Roger E Beaty, Yoed N Kenett, and James C Kaufman. 2024. The process definition of creativity. *Creativity Research Journal* 36, 3 (2024), 544–572.

[22] Zixun Guo, Jaeyong Kang, and Dorien Herremans. 2023. A domain-knowledge-inspired music embedding space and a novel attention mechanism for symbolic music modeling. In *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence (AAAI'23/IAAI'23/EAAI'23)*. AAAI Press, Article 566, 8 pages. https://doi.org/10.1609/aaai.v37i4.25635

[23] Matthew Guzdial, Nicholas Liao, Jonathan Chen, Shao-Yu Chen, Shukan Shah, Vishwa Shah, Joshua Reno, Gillian Smith, and Mark O Riedl. 2019. Friend, collaborator, student, manager: How design of an AI-driven game level editor affects creators. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*.

[24] Tellervo Härkki. 2023. Mobile gaze tracking and an extended linkography for collaborative sketching and designing. *International Journal of Technology and Design Education* 33, 2 (2023), 379–413.

[25] Gillian Hatcher, William Ion, Ross Maclachlan, Marion Marlow, Barbara Simpson, Nicky Wilson, and Andrew Wodehouse. 2018. Using linkography to compare creative methods for group ideation. *Design Studies* 58 (2018), 127–152.

[26] Tom Hewett, Mary Czerwinski, Michael Terry, Jay Nunamaker, Linda Candy, Bill Kules, and Elisabeth Sylvan. 2005. Creativity support tool evaluation methods and metrics. In *Creativity Support Tools: A workshop sponsored by the National Science Foundation*. 10–24.

[27] Qingqing Huang, Aren Jansen, Joonseok Lee, Ravi Ganti, Judith Yue Li, and Daniel P. W. Ellis. 2022. MuLan: A Joint Embedding of Music Audio and Natural Language. In *Proceedings of the 23rd International Society for Music Information Retrieval Conference, ISMIR 2022, Bengaluru, India, December 4-8, 2022*, Preeti Rao, Hema A. Murthy, Ajay Srinivasamurthy, Rachel M. Bittner, Rafael Caro Repetto, Masataka Goto, Xavier Serra, and Marius Miron (Eds.). 559–566. https://archives.ismir.net/ismir2022/paper/000067.pdf

[28] Youngseung Jeon, Seungwan Jin, Patrick C Shih, and Kyungsik Han. 2021. FashionQ: an AI-driven creativity support tool for facilitating ideation in fashion design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*.

[29] Jeff WT Kan and John S Gero. 2005. Can entropy indicate the richness of idea generation in team designing? In *CAADRIA05*. 451–457.

[30] Jeff WT Kan and John S Gero. 2008. Acquiring information from linkography in protocol studies of designing. *Design Studies* 29, 4 (2008), 315–337.

[31] Max Kreminski and John Joon Young Chung. 2024. Intent Elicitation in Mixed-Initiative Co-Creativity. In *Joint Proceedings of the ACM IUI Workshops*.

[32] Max Kreminski, Isaac Karth, Michael Mateas, and Noah Wardrip-Fruin. 2022. Evaluating mixed-initiative creative interfaces via expressive range coverage analysis. In *IUI Workshops*. 34–45.

[33] Max Kreminski and Michael Mateas. 2021. Reflective creators. In *International Conference on Computational Creativity*. 309–318.

[34] Arniel Labrada, Benjamin Bustos, and Ivan Sipiran. 2024. A convolutional architecture for 3D model embedding using image views. *The Visual Computer* 40, 3 (2024), 1601–1615.

[35] Tomas Lawton, Francisco J Ibarrola, Dan Ventura, and Kazjon Grace. 2023. Drawing with Reframer: Emergence and control in co-creative AI. In *Proceedings of the 28th International Conference on Intelligent User Interfaces*. 264–277.

[36] Mina Lee, Katy Ilonka Gero, John Joon Young Chung, Simon Buckingham Shum, Vipul Raheja, Hua Shen, Subhashini Venugopalan, Thiemo Wambsganss, David Zhou, Emad A. Alghamdi, Tal August, Avinash Bhat, Madiha Zahrah Choksi, Senjuti Dutta, Jin L.C. Guo, Md Naimul Hoque, Yewon Kim, Simon Knight, Seyed Parsa Neshaei, Antonette Shibani, Disha Shrivastava, Lila Shroff, Agnia Sergeyuk, Jessi Stark, Sarah Sterman, Sitong Wang, Antoine Bosselut, Daniel Buschek, Joseph Chee Chang, Sherol Chen, Max Kreminski, Joonsuk Park, Roy Pea, Eugenia Ha Rim Rho, Zejiang Shen, and Pao Siangliulue. 2024. A Design Space for Intelligent and Interactive Writing Assistants. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA)

(CHI '24). Association for Computing Machinery, New York, NY, USA.

[37] Seung Won Lee, Tae Hee Jo, Semin Jin, Jiin Choi, Kyungwon Yun, Sergio Bromberg, Seonghoon Ban, and Kyung Hoon Hyun. 2024. The Impact of Sketch-guided vs. Prompt-guided 3D Generative AIs on the Design Exploration Process. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*.

[38] Jialang Victor Li, Max Kreminski, Sean M Fernandes, Anya Osborne, Joshua McVeigh-Schultz, and Katherine Isbister. 2022. Conversation balance: A shared VR visualization to support turn-taking in meetings. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*.

[39] Antonios Liapis, Georgios Yannakakis, and Julian Togelius. 2013. Designer modeling for personalized game content creation tools. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, Vol. 9. 11–16.

[40] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J Cai. 2020. Novice-AI music co-creation via AI-steering tools for deep generative models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*.

[41] Ryan Louie, Jesse Engel, and Cheng-Zhi Anna Huang. 2022. Expressive communication: Evaluating developments in generative models and steering interfaces for music creation. In *Proceedings of the 27th International Conference on Intelligent User Interfaces*. 405–417.

[42] Todd I Lubart. 2001. Models of the creative process: Past, present and future. *Creativity Research Journal* 13, 3-4 (2001), 295–308.

[43] Shruti Mahajan, Leo Bunyea, Nathan Partlan, Dylan Schout, Casper Harteveld, Camillia Matuk, Will Althoff, Tyler Duke, Steven Sutherland, and Gillian Smith. 2019. Toward automated critique for student-created interactive narrative projects. In *Proceedings of the 15th AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-19)*.

[44] Atefeh Mahdavi Goloujeh, Anne Sullivan, and Brian Magerko. 2024. Is It AI or Is It Me? Understanding Users' Prompt Journey with Text-to-Image Generative AI Tools. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*.

[45] Jon McCormack, Camilo Cruz Gambardella, Nina Rajcic, Stephen James Krol, Maria Teresa Llano, and Meng Yang. 2023. Is writing prompts really making art? In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer, 196–211.

[46] Jon McCormack, Maria Teresa Llano, Stephen James Krol, and Nina Rajcic. 2024. No Longer Trending on Artstation: Prompt Analysis of Generative AI Art. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer, 279–295.

[47] Mark J Nelson, Swen E Gaudl, Simon Colton, and Sebastian Deterding. 2018. Curious users of casual creators. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*.

[48] Dmitry Nikolaev and Sebastian Padó. 2023. Representation biases in sentence transformers. *arXiv preprint arXiv:2301.13039* (2023).

[49] Peter Organisciak, Selcuk Acar, Denis Dumas, and Kelly Berthiaume. 2023. Beyond semantic distance: Automated scoring of divergent thinking greatly improves with large language models. *Thinking Skills and Creativity* 49 (2023), 101356.

[50] Lace Padilla, Matthew Kay, and Jessica Hullman. 2022. Uncertainty visualization. In *Computational Statistics in Data Science*. Wiley, 405–421.

[51] Morteza Pourmohamadi and John S Gero. 2011. LINKOgrapher: An analysis tool to study design protocols based on FBS coding scheme. In *DS 68-2: Proceedings of the 18th International Conference on Engineering Design (ICED 11)*.

[52] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning transferable visual models from natural language supervision. In *International Conference on Machine Learning*. PMLR, 8748–8763.

[53] 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

[54] Christian Remy, Lindsay MacDonald Vermeulen, Jonas Frich, Michael Mose Biskjaer, and Peter Dalsgaard. 2020. Evaluating creativity support tools in HCI research. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 457–476.

[55] Huda Salman. 2014. Linkography for evaluating ideas connectivity of Computer Aided Design-based protocols. In *eCAADe*, Vol. 1. 573–581.

[56] Téo Sanchez. 2023. Examining the Text-to-Image Community of Practice: Why and How do People Prompt Generative AIs? In *Proceedings of the 15th Conference on Creativity and Cognition*. 43–61.

[57] Ben Shneiderman. 2007. Creativity support tools: accelerating discovery and innovation. *Commun. ACM* 50, 12 (2007), 20–32.

[58] Amy Smith, Barrett R Anderson, Jasmine Tan Otto, Isaac Karth, Yuqian Sun, John Joon Young Chung, Melissa Roemmele, and Max Kreminski. 2025. Fuzzy Linkography: Automatic Graphical Summarization of Creative Activity Traces. *arXiv preprint arXiv:2502.04599* (2025).

[59] Gillian Smith and Jim Whitehead. 2010. Analyzing the expressive range of a level generator. In *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*.

[60] Saba Sturua, Isabelle Mohr, Mohammad Kalim Akram, Michael Günther, Bo Wang, Markus Krimmel, Feng Wang, Georgios Mastrapas, Andreas Koukounas, Andreas Koukounas, Nan Wang, and Han Xiao. 2024. jina-embeddings-v3: Multilingual Embeddings With Task LoRA. arXiv:2409.10173 [cs.CL] https://arxiv.org/abs/2409.10173

[61] Mikaela Angelina Uy, Jingwei Huang, Minhyuk Sung, Tolga Birdal, and Leonidas Guibas. 2020. Deformation-aware 3D model embedding and retrieval. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VII 16*. Springer, 397–413.

[62] Rosario Vidal, Elena Mulet, and Eliseo Gómez-Senent. 2004. Effectiveness of the means of expression in creative problem-solving in design groups. *Journal of Engineering Design* 15, 3 (2004), 285–298.

[63] Benjamin Warner, Antoine Chaffin, Benjamin Clavié, Orion Weller, Oskar Hallström, Said Taghadouini, Alexis Gallagher, Raja Biswas, Faisal Ladhak, Tom Aarsen, Nathan Cooper, Griffin Adams, Jeremy Howard, and Iacopo Poli. 2024. Smarter, Better, Faster, Longer: A Modern Bidirectional Encoder for Fast, Memory Efficient, and Long Context Finetuning and Inference. arXiv:2412.13663 [cs.CL] https://arxiv.org/abs/2412.13663

[64] Jason Wu, Yi-Hao Peng, Xin Yue Amanda Li, Amanda Swearngin, Jeffrey P Bigham, and Jeffrey Nichols. 2024. UIClip: a data-driven model for assessing user interface design. In *Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology*.

[65] Xiaohua Zhai, Basil Mustafa, Alexander Kolesnikov, and Lucas Beyer. 2023. Sigmoid loss for language image pre-training. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 11975–11986.

[66] Xiaoxuan Zhang, Zeping Zhan, Misha Holtz, and Adam M Smith. 2018. Crawling, indexing, and retrieving moments in videogames. In *Proceedings of the 13th International Conference on the Foundations of Digital Games*.

A Entropy Calculation on Fuzzy Linkographs

In traditional linkographs, forelink and backlink entropy are first computed on a move-by-move basis, by determining the probability that a link does or does not exist between a given move M and any following (forelinkable) or preceding (backlinkable) move; these values are then summed across all moves to give a total forelink and backlink entropy for the whole graph. **Horizonlink entropy** is a similar measure, but quantifies the unpredictability of links at each possible *horizon level*, i.e., each possible distance between pairs of moves; horizonlink entropy is first calculated for the set of all move pairs that are exactly *one* move apart, then for the set of all move pairs that are *two* moves apart, and so on, and these values are again summed together to give a total horizonlink entropy for the whole graph. Finally, the three different graph-level entropy values (forelink, backlink, and horizonlink) can themselves be summed to give an **overall link entropy** for the graph as a whole.

Entropy values can be calculated for fuzzy linkographs by treating the strength of a link between two moves (which already ranges from 0 to 1) as the *probability* of a binary link existing between these same moves. In Kan and Gero's original formulation of link entropy, $p(ON)$—the probability of a possible link existing—is calculated for each row of forelinks, backlinks, and horizonlinks (i.e., each "state" s) as the actual number of links that are present in this state divided by the maximum number of links n_s that are possible in this state. We thus calculate $p(ON)_s$ for the same states as:

$$p(ON)_s = \frac{\sum_{v \in L_s} v}{n_s} \tag{1}$$

...where L_s is the set of strength values v of all links that are present in state s. We then follow Kan and Gero's derivation [30] of $p(OFF)_s$ and all downstream entropy values.

One consequence of this interpretation is that a fuzzy linkograph in which every move is connected to every other by ambiguous links

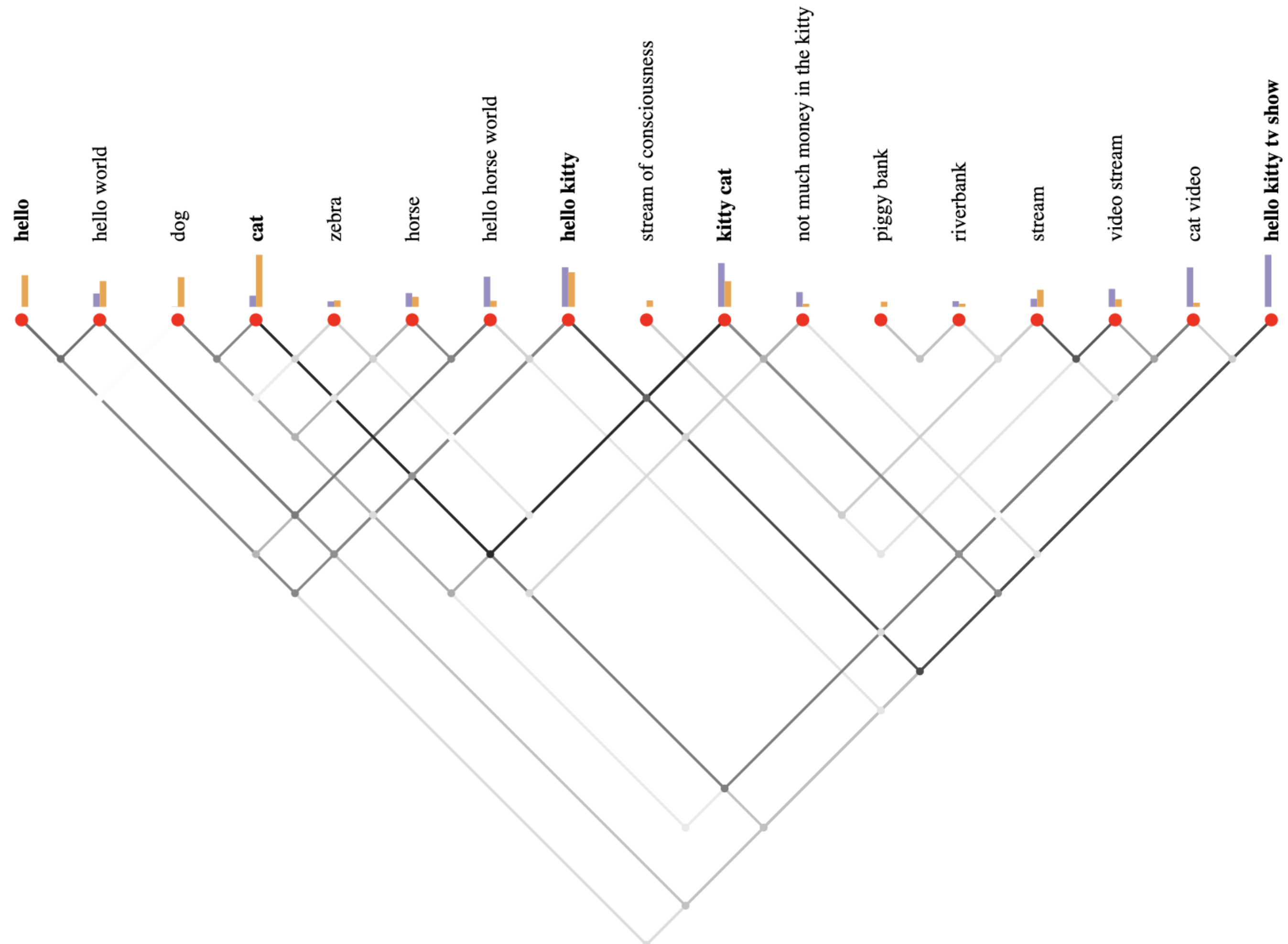

Figure 9: An example fuzzy linkograph of an open-ended stream-of-consciousness ideation activity. Links between moves are colored according to their strength (with darker lines indicating stronger links). Move text is displayed directly above the move markers. This graph features several critical moves: "hello" and "cat" are forelink critical moves that introduce the themes of greeting and cats respectively; "hello kitty" initially integrates these two themes; and "hello kitty tv show" is a backlink critical move, fully integrating these themes from the session's first half with the "video" theme from its second.

may be judged as having greater entropy than a linkograph with the same LDI, but an equal number of very strong and completely absent links. This is because the graph saturated with ambiguous links can be viewed as distributing uncertainty more evenly across the whole graph. Though we have not seen this cause any practical problems so far, it strikes us as somewhat counterintuitive. At any rate, we do not make evaluative use of entropy in the analyses we present here, because our goal is not to comparatively evaluate the success of different creative practices or tools but to broadly characterize creative behaviors observed in different situations; further refinement of entropy measures on fuzzy linkographs may be a good topic for future work.

B Limitations and Future Work

A fuzzy linkograph is fuzzy: it does not perfectly match the linkograph that a human annotator might produce for the same design episode. This is partly because even human annotators tend to disagree to some extent about how to construct linkographs from the same design moves [25], but also partly because embedding models are imperfect proxies for human perceptions of design move relatedness. In particular, transformer-based sentence embedding models tend to place disproportionate emphasis on certain parts of speech [48] and may overlook less obvious semantic relationships between design moves in some cases; for instance, a link was not inferred between "not much money in the kitty" and "piggy bank" in Figure 9, even though both involve currency.

Figure 10: *Linkographic abundance*: fuzzy linkographs of 100 randomly selected image prompting traces, laid out as thumbnails. As new activity traces are logged, CST researchers might automatically generate linkographs for these traces and monitor the overall distribution, or use interesting-looking linkographs as jumping-off points for deeper investigation.

To construct our linkographs, we used `all-MiniLM-L6-v2` [53]—an open, general-purpose embedding model that has previously been validated against a human baseline for assessment of semantic similarity in creativity research [2]. However, better performance could likely be achieved with an embedding model specifically tuned on data from human-constructed linkographs. It may also be possible to use a large language model (or some alternative approach) rather than an embedding model to predict links between design moves; some automated creativity assessment pipelines seem to perform better when embeddings are replaced with an LLM that has been fine-tuned on task-specific data [49], though this would likely increase cost per linkograph considerably.

Our reformulation of link entropy measures for fuzzy linkography is not necessarily the best one possible: in particular, by averaging continuous link strengths and treating the result as a probability of a binary link either existing or not existing, we arguably conflate association strength and probability of association in a way that could sometimes disguise meaningful differences between linkographs. To some extent this is an inherent risk of flattening complex graph structures down to summary values, and our entropy results on real data still seem to follow intuitive expectations—but it may nevertheless be desirable to invent an alternative quantitative proxy for design episode "dynamism" that yields comparatively higher values for fuzzy linkographs consisting of mixed weak and strong links. Similar criticisms have previously been leveled of the conventional approach to calculating link entropy in general; alternative formulations of entropy proposed in the past [9, 10] may be adaptable to fuzzy linkography.

Linkography itself (especially when viewed as a visualization technique) does not necessarily scale well with trace length past a certain threshold: a linkograph containing more than a few dozen moves tends to become visually overwhelming. However, clustering or summarization of moves may permit condensation of longer traces to effectively shorter ones for linkographic purposes [62], and quantitative linkographic metrics may still support trace clustering, critical move identification, and so on even at larger scales.