

Refining the Computational Analysis of Interactive Fiction: A Hybrid Framework using Bag-of-Paths Formalism and LLMs

Extended Abstract

Guillaume Guex

Department of Language and Information Sciences, University of Lausanne
guillaume.guex@unil.ch

The computational study of *interactive fiction* (Aarseth, 1997; Montfort, 2005) generally involves extracting a directed graph where nodes represent narrative units (e.g. paragraphs in gamebooks) and edges represent the choices offered to the player (Riedl and Bulitko, 2013; Métrailler, 2025). From there, mathematical tools can be applied to understand the global narrative structure. For instance, the *Bag-of-Paths formalism* (Françoisse et al., 2017; Guex et al., 2021) allows us to define a probability distribution over all possible paths, corresponding to unique storylines, based on a "temperature" parameter (balancing between random choices and efficiency). This approach enables the computation of exact metrics in closed form, such as the probability of encountering a specific paragraph, the correlation between node presences (indicating linked narrative events), and the most probable paths through the entire story.

However, the Bag-of-Paths formalism deals strictly with the *topology* or the *structure* of the book. It treats the work as a skeleton, missing one of the most important components of interactive fiction: the *narrative content*. Consequently, three aspects remain difficult to analyze using this mathematical approach alone: (1) the nature of the transition (is the choice difficult, dangerous, or evident?); (2) the narrative relationship between correlated paragraphs; and (3) the global narrative resulting from a particular path (does the resulting story make sense?).

The current submission proposes to refine the Bag-of-Paths approach using Large Language Models (LLMs) to address these three cases. Indeed, the large volume of textual elements to analyze, comprising local choices, node pairs, and thousands of potential generated stories, makes a comprehensive close reading labor-intensive and time-consuming, even for a single book. To overcome this scalabil-

ity issue, we automate the analysis using a pipeline of structured prompts designed to enforce specific output formats (e.g., JSON), ensuring that semantic features are extracted consistently for the mathematical model (Wang et al., 2025). We therefore employ LLMs at the three levels described above as follows:

1. Instead of assuming uniform probabilities at each node, we use LLMs to analyze choice texts and characterize transitions (e.g., danger level, moral implications). These scores can be used to build various transition probabilities within the Bag-of-Paths formalism, allowing us to simulate specific playstyles (e.g., prudent vs. reckless).
2. We compare the structural correlation of nodes with their semantical similarity computed by LLMs. This allows us to detect "narrative coherence" or "dissonance", identifying instances where the graph forces a connection between semantically related or unrelated events.
3. Finally, we generate a set of representative paths according to the probability model and use the LLMs as an "AI critic" to assess the story as a whole. This step moves beyond local transition analysis to construct global indices, evaluating whether the generated stories actually yield a coherent and compelling narrative experience.

We apply this framework to the first books of the *Lone Wolf series* by Joe Dever (Dever and Chalk, 1991), using the open data provided by Project Aon¹. We demonstrate that this hybrid approach allows constructing new indices which capture, for instance, the difficulty of the adventure, the feeling of freedom, or narrative coherence better than

¹<https://www.projectaon.org>

standard graph metrics. The long-term goal is to establish a robust framework for the "distant reading" of branching narratives, bridging the gap between topology and semantics.

References

- Espen J. Aarseth. 1997. *Cybertext: perspectives on ergodic literature*. Johns Hopkins University Press, Baltimore, Md. London.
- Joe Dever and Gary Chalk. 1991. *Flight from the dark*. Red Fox, London. OCLC: 226242649.
- Kevin Françoisse, Ilkka Kivimäki, Amin Mantrach, Fabrice Rossi, and Marco Saerens. 2017. [A bag-of-paths framework for network data analysis](#). *Neural Networks*, 90:90–111.
- Guillaume Guex, Sylvain Courtain, and Marco Saerens. 2021. [Covariance and correlation measures on a graph in a generalized bag-of-paths formalism](#). *Journal of Complex Networks*, 8(6):cnaa026.
- Nick Montfort. 2005. *Twisty little passages: an approach to interactive fiction*. MIT Press, Cambridge, Mass. London.
- Coline Eloïse Métrailler. 2025. *Une étude plurielle des réseaux de personnages : narrations interactives, jeux vidéo et adaptations*. Ph.D. thesis.
- Mark O. Riedl and Vadim Bulitko. 2013. [Interactive Narrative: An Intelligent Systems Approach](#). *AI Magazine*, 34(1):67–77.
- Darren Yow-Bang Wang, Zhengyuan Shen, Soumya Smruti Mishra, Zhichao Xu, Yifei Teng, and Haibo Ding. 2025. [SLOT: Structuring the Output of Large Language Models](#). *arXiv preprint*. Version Number: 1.