

Treating Games as Plays? Computational Approaches to the Detection of Scenes in Game Dialogs

Martin Schlenk^{1,*†}, Thomas Efer^{2,*†} and Manuel Burghardt^{3,*†}

¹Computational Humanities, Leipzig University, Augustusplatz 10, Leipzig, 04109, Germany

²Computational Humanities, Leipzig University, Augustusplatz 10, Leipzig, 04109, Germany

³Computational Humanities, Leipzig University, Augustusplatz 10, Leipzig, 04109, Germany

Abstract

Digital games are a complex multimodal phenomenon that is examined in a variety of ways by the highly interdisciplinary field of game studies. In this article, we focus on the structural aspect of the diegetic language of games and examine the extent to which established methods of computational drama analysis can also be successfully applied to digital games. Initial experiments show that both games and drama texts have an inventory of characters that drive the plot forward. In dramas, this plot is usually subdivided into individual acts and scenes. In games, however, such systematic segmentation is the exception rather than the rule, or if it is present, it is implemented very differently in different games. In this paper, we therefore focus on exploring alternative ways of making scene-like structures in game dialogs identifiable with the help of computers. As a result of these experiments, exciting future perspectives emerge that raise the question of whether computer-aided methods of scene recognition, which are inspired by media such as games and films, can also be applied to classical dramas in the future in order to fundamentally question their historical-editorial scene classification.

Keywords

computational game studies, drama analysis, scene detection, game script analysis, dialog sequence analysis

1. Introduction to Computational Drama Analysis

Game studies has emerged as a significant area of research, recognizing digital games as highly interactive and multimodal cultural artifacts.¹ This complexity presents considerable challenges for the large-scale analysis of digital games, leading to a predominance of qualitative and hermeneutic research within the field.² While computational humanities methods have been applied to various aspects of games [2], these efforts have typically focused on individual

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark

*Corresponding author.

†These authors contributed equally.

✉ martin.schlenk@informatik.uni-leipzig.de (M. Schlenk); efer@informatik.uni-leipzig.de (T. Efer);

burghardt@informatik.uni-leipzig.de (M. Burghardt)

🌐 <https://www.mathcs.uni-leipzig.de/en/ifi/research/computational-humanities#c685628> (M. Schlenk);

<https://www.uni-leipzig.de/en/profile/mitarbeiter/dr-thomas-efer> (T. Efer);

<https://www.uni-leipzig.de/en/profile/mitarbeiter/juniorprof-dr-manuel-burghardt> (M. Burghardt)

🆔 0009-0006-3125-2405 (M. Schlenk); 0000-0002-8376-3884 (T. Efer); 0000-0003-1354-9089 (M. Burghardt)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹There are various overviews and introductions to game studies, for example [10] and [7].

²This is reflected by most of the articles of “The international journal of computer game research” (<https://gamestudies.org/>), which has been the major venue for game studies since 2001.

modalities, such as visual aspects, game music or game language. Among these, the “language of gaming” [3] is probably the best researched modality. It can be categorized into *orthogame* language, which encompasses elements like menus and in-game dialogs, and *paragame* language, referring to language used outside the game context, such as in reviews and forums [3]. By focusing on orthogame language, particularly character dialogs, numerous parallels to the literary genre of stage plays become evident. If we conceptualize games as plays, game characters can be seen as actors performing dialogs on the virtual stage of a diegetic game world. This analogy, treating digital games as a form of stage play, opens up a wide range of computational approaches previously explored in the computational humanities under the term *drametrics* [13, 12] or more recently as computational drama analysis [1].

The computational drama analysis toolbox has produced a number of interesting approaches in recent years. Inspired by Solomon Marcus’ “Mathematical Poetics” (1970) [6], Wilhelm, Burghardt, and Wolff [16] as well as Schmidt, Burghardt, Dennerlein, and Wolff [13] have presented tools that analyze and visualize drama texts according to Marcus’ proposed configuration matrices, which is basically a table with occurrences of characters in different scenes of a drama. As the notion of configuration matrices is closely related to the idea of an adjacency matrix, which in turn can be translated into graph structures immediately, it is not surprising that another branch of computational drama analysis has investigated to model and analyze drama texts as graphs or social networks (see [8, 14, 15, 5]) which has ultimately lead to a dedicated infrastructure for computational research on drama texts: ³[4].

By leveraging these approaches for computational drama analysis to the area of digital games, researchers can gain novel insights into the systematic and structural elements of game narratives and dialogs. The proposed approach is also heavily inspired by the availability of a novel corpus resource, the “Video Game Dialogue Corpus” (VGDC)⁴, which provides access to the dialogs of more than 50 video games with a total of 6.2 million words and over 13,000 different character labels.

2. Toward Dialog-Based Scene Detection in Video Games

The above related works make clear that computational drama analysis heavily relies on character information (*who speaks?*) but also on structural information, such as acts and scenes (*who is on stage together?*). While intuitively we might assume similar structural units in digital games, for instance levels or acts in a game world, this actually is not the case for most games. The majority of games does not have a systematic segmentation of the plot, and those that have segments do it in largely different ways that cannot easily be compared to one another.

In order to make the methods known from computational drama analysis available for the large-scale analysis of game dialog, we explore ways of dialog-based scene detection in video games. This work in progress adds to existing research that has been conducted for scene detection in narrative texts, such as novels and biographies [17], but also for dramatic texts [11]. While these existing works heavily rely on transformer-based experiments, we present a robust algorithm that is inspired by the work of Nalisnick and Baird [9], as it utilizes character

³Drama Corpora Project: <https://dracor.org/>

⁴<https://correlation-machine.com/VideoGameDialogueCorpus/>

interactions.

Our approach identifies significant changes in the composition of speaking characters within the game narrative as key indicators of scene transitions. We analyze dialogue sections to create segments based on the relative homogeneity of character interactions. This method allows us to examine the game’s narrative structure in a way that is analogous to the act and scene divisions in traditional drama⁵. Alternative segmentation approaches could have been employed, such as dividing the text based on character appearances, using topic modeling to identify thematic shifts, or applying semantic analysis to determine when the subject of conversation changes. Each of these methods would offer different insights into the games’ narrative structures. Our choice of method was guided by the aim of creating segments that most closely resemble the function of acts and scenes in traditional drama.

3. Game Corpus

For our first experiments we rely on the above mentioned “Video Game Dialogue Corpus” (VGDC), which consists of JSON-formatted files containing transcripts of dialogs from various games. These files form the core of our analysis, providing a rich dataset of in-game conversations, actions, and narrative events. Additionally, we have access to metadata files that offer supplementary information, such as character aliases, which aid in the accurate identification and analysis of speakers throughout the games.

Throughout the corpus, character dialogue is the only consistently available structural element. This focus on character interactions enables standardized analysis across different games, despite variations in narrative structure or gameplay mechanics.

The transcript files are structured to capture the spoken dialogue, which is crucial for understanding the narrative flow. This structure allows us to analyze the verbal elements of the games’ storytelling, providing insights into character development and plot progression.

From the more than 50 available games in the corpus, we picked *Final Fantasy VI* for the upcoming examples. Released in 1994, *Final Fantasy VI* is noted for its large cast of characters and its narrative complexity. It is part of the *Final Fantasy* series, which consists of Japanese fantasy role-playing games where players control a group of characters and experience an adventure from their perspective. These games are known for their complex narratives and character development.

4. Scene-Detection in Game Dialogs

Our approach to detecting scene-like structures in video game dialogs is based on analyzing character interactions and their dynamics throughout the game script. Central to this approach are the concepts of activity and inactivity thresholds. The activity threshold determines the relative proportion of dialogue lines at which a character is considered “active” in a scene. A low threshold results in characters being classified as active more quickly, typically leading

⁵As players, we can often intuitively recognize scene changes or major plot points while playing the game. However, these transitions are only scarcely explicitly marked in the dialogue transcripts, necessitating our computational segmentation process.

to longer scenes. Conversely, a high threshold results in shorter scenes, as characters must contribute more dialogue to be considered active. The inactivity threshold defines how long a character can remain silent before being considered inactive. A low threshold leads to more frequent scene changes, as characters are more quickly considered inactive. A high threshold allows for longer periods of inactivity, resulting in more stable and longer scenes. These thresholds are dynamically adjusted based on the number of active characters in a scene. This enables flexible adaptation to various dialogue situations, from intimate one-on-one conversations to complex group interactions. After the initial scene detection, we perform a post-processing step where adjacent scenes with high similarities are merged. This step serves to optimize very small scenes and achieve a more coherent overall structure.

4.1. Technical Implementation

To operationalize these concepts, we developed a Python-based analytical tool. The technical implementation of our method comprises several steps:

Data preprocessing: We begin by loading and processing the dialogue data from the JSON file for Final Fantasy VI. This data contains structured information about character dialogs and metadata. **Character name normalization:** To ensure consistent identification of characters throughout the analysis, we implement a normalization process. This process accounts for various aliases and name variations that may occur in the game script. For example, a character might be referred to by a nickname, full name, or title at different points in the story. Our normalization function uses the metadata to map these variations to a standard identifier for each character, improving the accuracy of our scene detection algorithm. **Dynamic threshold calculation:** Our algorithm adjusts the activity and inactivity thresholds based on the number of characters in a scene. As the number of characters increases, the activity threshold decreases, reflecting the expectation that individual characters speak less in crowded scenes. Conversely, the inactivity threshold increases, allowing for longer periods of silence before a character is considered inactive in larger group settings.

The calculation of the dynamic thresholds for activity (t_A) and inactivity (t_I) is based on the following formulas:

$$t_A = \max\left(t_{A_{base}} - (N - 2) \times \frac{1}{20}, 0\right) \quad (1)$$

$$t_I = \min\left(t_{I_{base}} + (N - 2) \times \frac{1}{50}, 1\right) \quad (2)$$

Where the base thresholds $t_{A_{base}}$ and $t_{I_{base}}$ are distinct predefined values that can be independently adjusted in the interactive analysis environment. The value N represents the current number of active characters in a scene. The fractions $\frac{1}{20}$ and $\frac{1}{50}$ are empirically determined adjustment factors that modify the thresholds based on the number of characters, chosen for their interpretability and smooth scaling properties.

The subtraction of 2 from N serves as a reference point, assuming a baseline scenario typically involves two characters. From this baseline, the thresholds are adjusted as the number of characters increases or decreases.

In our analysis, we iterate over a range of base threshold values to explore how different settings affect the scene detection. The activity threshold decreases as the number of characters increases, allowing for more inclusive scene participation in larger groups. Conversely, the inactivity threshold increases with more characters, accommodating longer periods of individual silence in group settings.

These thresholds are not scores but direct boundaries determining when a character is considered active or inactive in a scene. A character is deemed active if their dialogue proportion exceeds the activity threshold, and inactive if their silence exceeds the inactivity threshold.

These dynamic adjustments enable our system to adapt flexibly to various scenarios, from intimate dialogs between two individuals to complex group interactions involving multiple participants, thereby providing a nuanced analysis of the game's dialogue structure. The use of separate base thresholds for activity and inactivity, each with its own slider in the interactive tool, allows for fine-tuned control over the analysis, accommodating different aspects of dialogue pacing and character involvement.

Scene analysis: The core of our method involves processing the dialogue data sequentially, identifying scene boundaries based on character activity patterns. This includes tracking active characters and their dialogue contributions, applying dynamic thresholds, and detecting significant changes in the active character set that indicate potential scene transitions.

Our implementation includes an interactive analysis environment that allows for real-time adjustment of parameters and immediate visualization of results. This environment, as shown in Figure 1, provides sliders for adjusting the activity threshold, inactivity threshold, and merge threshold.

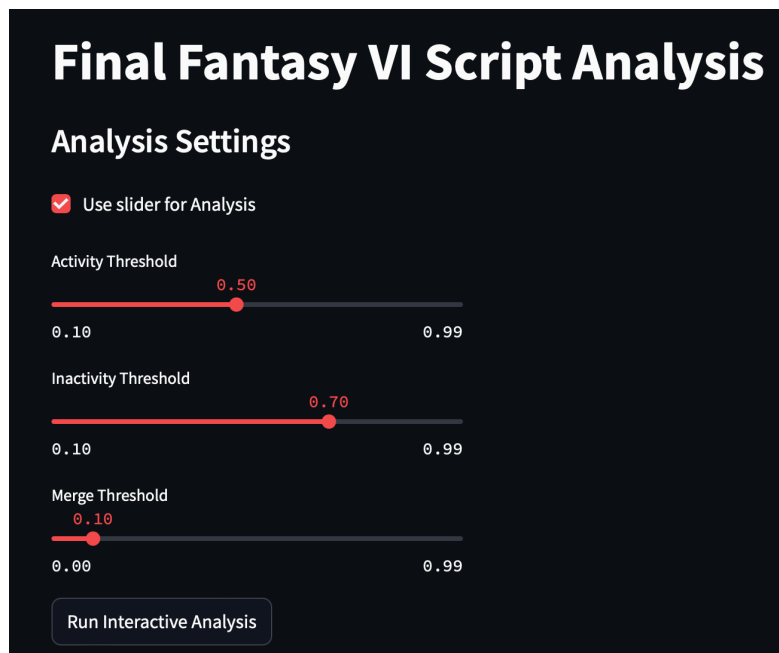


Figure 1: Interactive analysis environment for Final Fantasy VI Script

This interactive approach enables researchers to experiment with different threshold combinations and observe their impact on scene detection. Users can fine-tune the analysis parameters to best suit the specific narrative structure of Final Fantasy VI, allowing for a more nuanced understanding of the game’s dialogue dynamics and character interactions. The ability to adjust these parameters in real-time and visualize the results immediately provides a powerful tool for exploring the narrative structure of the game. It allows for rapid iteration and hypothesis testing, facilitating a more thorough and insightful analysis of the game’s script.

It’s important to note that dialogue and scene lengths significantly impact our thresholding approach. Longer dialogues may require different thresholds compared to shorter exchanges. To address this, our method uses adaptive thresholding based on the game’s overall dialogue density, adjusting to different narrative pacing and styles across games or within game sections.

5. Results and Discussion

Our analysis of the Final Fantasy VI script using our scene detection algorithm yielded intriguing insights into the narrative composition of the game. We conducted multiple analyses with varying thresholds to observe how these parameters affect scene delineation.

5.1. Combined Scene Structure Analysis

Figure 2 presents a visualization of how our scene segmentation changes as we adjust the base inactivity threshold while keeping the merge threshold (0.96) and base activity threshold (0.7) constant.

This visualization reveals several key findings:

1. **Threshold sensitivity:** At lower inactivity thresholds (0.45-0.55), we observe numerous short “scenes”, indicating over-segmentation. As the threshold increases (0.85-0.97), we see fewer, longer segments, potentially under-segmenting the narrative. The middle range (approximately 0.65-0.80) appears to produce a more balanced segmentation with scene lengths that intuitively align with narrative structures.
2. **Local stability:** Notably, certain segment boundaries remain consistent across multiple threshold values. This is visually evident where vertical black lines align across different threshold levels, particularly around dialogue line numbers 800, 1680, but also at several other points. This local stability suggests that these boundaries likely correspond to significant shifts in speaker composition or narrative focus.
3. **Analytical robustness:** The presence of these stable boundaries across varying thresholds underscores the robustness of our analysis tool. It indicates that our algorithm consistently identifies major narrative transitions, even as we adjust the sensitivity of our scene detection parameters.

The ability to observe these patterns was greatly facilitated by our exploratory tool, which allowed for rapid experimentation with different threshold values. This interactive approach enabled us to identify optimal ranges for scene segmentation and to recognize the persistence of certain narrative breakpoints across various analytical configurations. These findings not

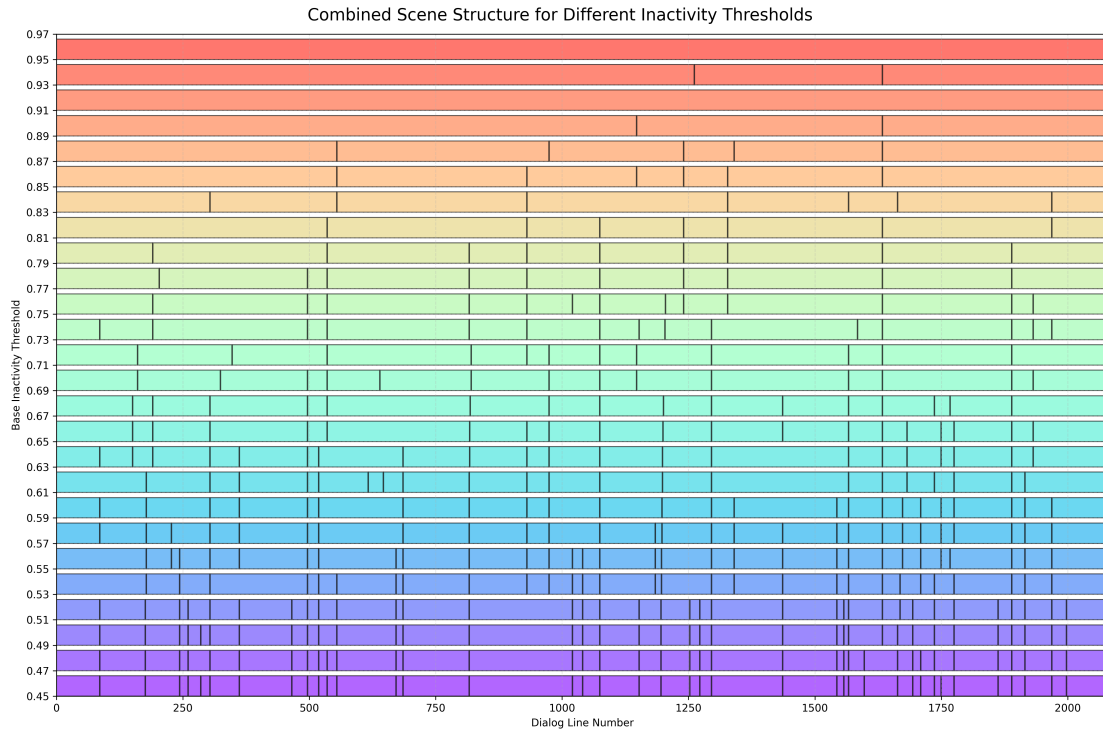


Figure 2: Combined Scene Structure for Different Inactivity Thresholds

only provide insights into the structure of Final Fantasy VI’s narrative but also demonstrate the value of our flexible, threshold-based approach to scene detection in game dialogs. The balance between sensitivity to narrative changes and the avoidance of over-segmentation appears achievable within a specific range of threshold values, offering a promising framework for analyzing complex game narratives.

5.2. Character Dialogue Distribution Analysis

Building upon our scene structure analysis, we now turn our attention to the distribution of character dialogs within these detected scenes. This analysis provides insights into the narrative focus and character prominence throughout the game.

Figure 3 presents the dialogue distribution for the eight most vocal characters in Final Fantasy VI, with $t_{A_{base}}$ of 0.70 and $t_{I_{base}}$ of 0.87.

This visualization reveals several large segments, which could be loosely compared to acts in a traditional drama, though such comparisons should be made cautiously. Each segment shows varying proportions of dialogue from different characters, indicating shifts in narrative focus. It’s important to note that while only the eight most vocal characters are represented, other characters may also contribute to the dialogue in these scenes. When we adjust the inactivity threshold to 0.71 (Figure 4), we observe a fragmentation of these larger segments into smaller scenes. This fragmentation reveals more granular patterns in character interactions.

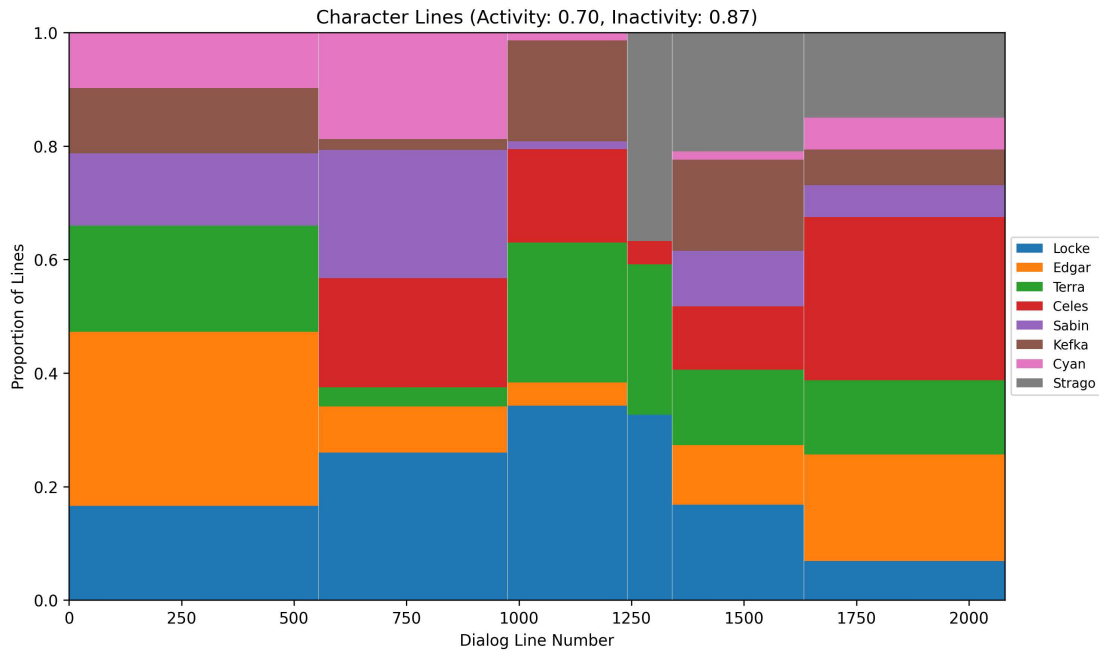


Figure 3: Character Lines Distribution (Activity: 0.70, Inactivity: 0.87)

Notably, not all of the eight most vocal characters appear in every scene, providing insights into the story’s structure and character groupings throughout the narrative. Further reducing the inactivity threshold to 0.57 (Figure 5) results in even shorter scenes, including instances where none of the eight most vocal characters speak (represented by white lines). The appearance of these “silent” scenes only at very low inactivity thresholds suggests that the eight most vocal characters are well-distributed throughout the game’s narrative. This distribution indicates a balanced approach to character involvement in Final Fantasy VI’s storytelling.

These visualizations demonstrate the flexibility of our analysis tool in capturing different levels of narrative granularity. By adjusting thresholds, we can explore the story’s structure from broad narrative arcs to more detailed character interactions, providing a multi-faceted view of the game’s dialogue composition.

5.3. Limitations and Future Directions

While our analysis provides valuable insights into the dialog structure of Final Fantasy VI, several limitations and opportunities for future research should be noted: **Corpus utilization:** Our current study focuses on a single game from the Video Game Dialogue Corpus (VGDC). The next step would be to extend our analysis to other games within this corpus, enabling more comprehensive and comparative analyses across a wider range of titles, series, and genres. This expansion would significantly enhance the generalizability of our findings and provide a broader perspective on dialog structures in video games.

Method expansion: Our scene detection method currently relies primarily on dialogue

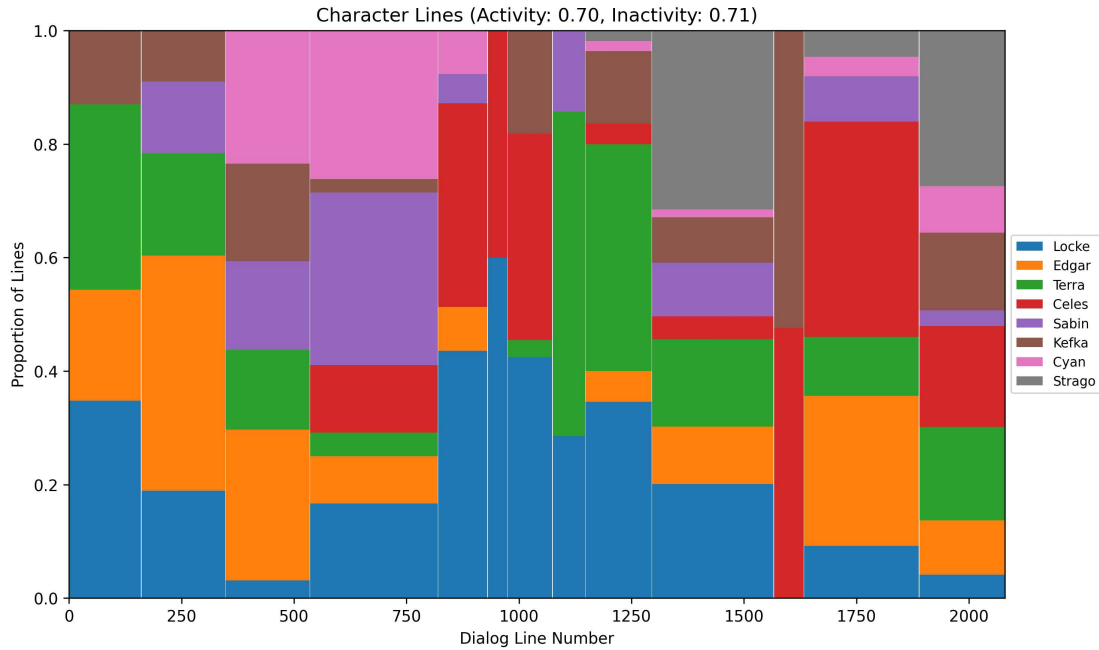


Figure 4: Character Lines Distribution (Activity: 0.70, Inactivity: 0.71)

data. Future iterations could incorporate additional information present in the VGDC transcript files, such as action descriptions or location changes, to create a more nuanced understanding of scene boundaries. This expansion would leverage the full potential of the available data and potentially improve the accuracy of scene detection.

Game linearity: The relatively linear dialog structure of Final Fantasy VI was advantageous for our current analytical approach. Future work should aim to adapt our approach for less linear games, developing solutions that can handle branching narratives and player-driven story progression. This adaptation would broaden the applicability of our method to a wider range of game structures.

Integration with other approaches: Comparing our dialogue-based scene detection method with other segmentation approaches, such as those based on visual or audio cues, could provide a more holistic understanding of dialog structures in video games. This integration could lead to more robust and comprehensive analytical tools for game narrative analysis.

These directions for future research aim to address the current limitations of our study while expanding the scope and applicability of computational methods in game narrative analysis. By continuing to develop and refine these approaches, we can deepen our understanding of storytelling techniques in video games and their relationship to other narrative media.

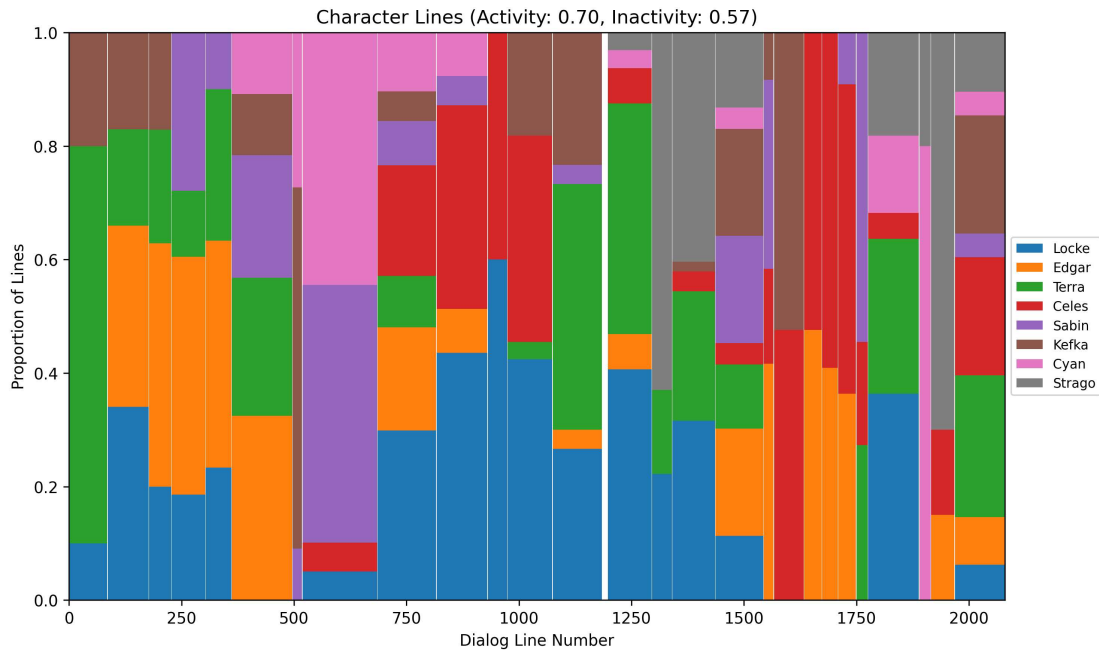


Figure 5: Character Lines Distribution (Activity: 0.70, Inactivity: 0.57)

References

- [1] M. Andresen and N. Reiter, eds. *Computational Drama Analysis*. DeGruyter, 2024. DOI: 10.1515/9783111071824.
- [2] M. Burghardt and V. Piontkowitz. “Computational Game Studies? Drei Annäherungsperspektiven”. In: *Proceedings of the 10th Conference of the German association of Digital Humanities “Digital Humanities im deutschsprachigen Raum – DHd”*. Passau, 2024. DOI: 10.5281/zenodo.10698401.
- [3] A. Ensslin. *The Language of Gaming*. 2011. DOI: 10.1007/978-0-230-35708-2.
- [4] F. Fischer, I. Börner, M. Göbel, A. Hechtel, C. Kittel, C. Milling, and P. Trilcke. “Programmable Corpora: Introducing DraCor, an Infrastructure for the Research on European Drama”. In: *Proceedings of the Digital Humanities Conference*. 2019. DOI: 10.5281/zenodo.4284002.
- [5] F. Fischer, M. Göbel, D. Kampkaspar, C. Kittel, and P. Trilcke. “Network Dynamics, Plot Analysis: Approaching the Progressive Structuration of Literary Texts”. In: *Proceedings of the Digital Humanities Conference*. 2017.
- [6] S. Marcus. *Poetica Matematica*. Bucharest: Editura Academiei Republicii Socialiste Romania, 1970.
- [7] F. Mäyrä. *An Introduction to Game Studies*. Sage, 2008.

- [8] F. Moretti. “Network Theory, Plot Analysis”. In: Stanford Literary Lab Pamphlets, 2011. Chap. 2.
- [9] E. T. Nalisnick and H. S. Baird. “Character-to-Character Sentiment Analysis in Shakespeare’s Plays”. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Sofia, Bulgaria, 2013, pp. 479–483.
- [10] D. B. Nieborg and J. Hermes. “What is game studies anyway?” In: *European Journal of Cultural Studies* (2008), pp. 131–146. doi: 10.1177/1567549407088328.
- [11] J. Pagel, N. Sihag, and N. Reiter. “Predicting Structural Elements in German Drama”. In: *Proceedings of the Second Conference on Computational Humanities Research*. 2021.
- [12] M. Romanska. “Drametrics: what dramaturgs should learn from mathematicians”. In: *The Routledge companion to dramaturgy*. Routledge, 2014, pp. 438–447.
- [13] T. Schmidt, M. Burghardt, K. Dennerlein, and C. Wolff. “Katharsis – A Tool for Computational Drametrics”. In: *Proceedings of the Digital Humanities Conference*. Utrecht, 2019.
- [14] P. Trilcke. “Social Network Analysis (SNA) als Methode einer textempirischen Literaturwissenschaft.” In: Münster: Mentis, 2013, pp. 201–247.
- [15] P. Trilcke, F. Fischer, and D. Kampkaspar. “Digital network analysis of dramatic texts”. In: *Proceedings of the Digital Humanities Conference*. 2015.
- [16] T. Wilhelm, M. Burghardt, and C. Wolff. ““To See or Not to See” – An Interactive Tool for the Visualization and Analysis of Shakespeare Plays”. In: (2013).
- [17] A. Zehe, L. Konle, L. Dümpelmann, E. Gius, A. Hotho, F. Jannidis, L. Kaufmann, M. Krug, F. Puppe, N. Reiter, A. Schreiber, and N. Wiedmer. “Detecting Scenes in Fiction: A new Segmentation Task”. In: 2021, pp. 3167–3177. doi: 10.18653/v1/2021.eacl-main.276.