Helen Gao

HUM346 Research Proposal

Comparing TV Adaptations of *The Three-Body Problem*

Narrative and Research Question

The Three-Body Problem, the first book in the *Remembrance of Earth's Past* trilogy, is a Chinese science fiction novel written by Liu Cixin in 2008 and translated into English by Ken Li in 2014. The story covers several decades, from the time of China's Cultural Revolution to the modern era, and it revolves around humanity's encounter with an advanced alien civilization. The novel also incorporates complex concepts from the relevant scientific fields, and it is "full of lovingly lengthy passages of technical exposition about everything from quantum mechanics to artificial intelligence"[1]. Recently, this book has been adapted into two separate television series: one produced by the Chinese streaming site Tencent Video, and the other released by American streaming giant Netflix.

The Tencent adaptation, titled *Three-Body,* premiered in January 2023. *Three-Body* is a 30-episode adaptation of the trilogy, and it has been praised for "its faithfulness, in broad outline, to the trilogy's first novel, 'The Three-Body Problem,' on which it is based"[2]. Faithfulness to the original is often important to fans, but a critic comments that "For a general audience, the show's similarities to the book may be more problematic"[3]. Additionally, given the original novel's

[1] Jason Heller, "'Three-Body Problem' Asks A Classic Sci-Fi Question, In Chinese," *NPR*, November 13, 2014, sec. Book Reviews, https://www.npr.org/2014/11/13/363123510/three-body-problem-asks-a-classic-sci-fi-question-in-chinese.
[2] Mike Hale, "'Three-Body' Review: A Chinese Series Beats Netflix to the Screen," *The New York Times*, February 3, 2023, sec. Arts, https://www.nytimes.com/2023/02/03/arts/television/three-body-review.html.
[3] Hale, "'Three-Body' Review."

dedication to explaining scientific topics, this adaptation could feel inaccessible to non-technical audiences.

The Netflix adaptation, titled *3 Body Problem*, was released in March 2024. It was created by the same people who made *Game of Thrones*, and it takes many more creative liberties with regard to the plot. For instance, the modern parts of the story are moved from Beijing to London and Oxford. Additionally, while *Three-Body* doesn't entirely follow the characters of the original book – "contemporary characters, mostly female, are added or greatly expanded upon"[4] – the creators of *3 Body Problem* do away with the main character of the book entirely, and "shuffle and reconfigure characters from throughout Liu's trilogy into a clique of five attractive Oxford-grad prodigies who carry much of the narrative"[5]. *3 Body Problem* is also described as having "given Liu's heavy science a dose of the humanities"[6], allowing the characters to speak and interact in more relatable ways.

It is relatively uncommon for a book to have both Chinese and American TV adaptations, especially ones released within two years of each other; typically, if a TV adaptation already exists, streaming services would simply choose to translate and subtitle it for a localized release. Thus, studying both *Three-Body* and *3 Body Problem* as adaptations of the same source material into the same medium can provide insights into how creators from different cultures, writing for different audiences, may balance scientific explanations with other plot points, explore the existential themes of a story about alien contact, and develop or reinterpret characters to fit certain perspectives. These choices can reflect cultural differences, political sensitivities, and

[4] Hale, "'Three-Body' Review."
[5] James Poniewozik, "Review: '3 Body Problem' Is a Galaxy-Brained Spectacle," *The New York Times*, March 20, 2024, sec. Arts, https://www.nytimes.com/2024/03/20/arts/television/3-body-problem-netflix-review.html.
[6] Poniewozik, "Review."

specific storytelling techniques. Thus, a research question emerges: How do the narrative and thematic elements of Tencent and Netflix's adaptations of The Three-Body Problem differ?

Dataset Identification

For this project, the dataset would consist of the transcripts of every episode in each series. Though other elements such as musical scores, casting choices, and set design also contribute greatly to the experience of watching a TV show, using transcripts helps ensure that the dataset is objectively accurate, and it enables textual analysis. Moreover, there is precedent for analyzing TV series in this manner; for example, a project that examines differences between episodes of the TV series *Doctor Who* based on the showrunner uses a "corpus of transcriptions" to conduct a stylometric analysis.

For the purposes of comparison, this analysis would require transcripts of *Three-Body*, which is Chinese, to be translated into English. Luckily, these translations are available online as text files on Google Drive, as part of the Sinodex project (though unfortunately, the site for this project is no longer reachable). Though the files are labeled as transcripts, their contents appear to be in the format of subtitles. Below are the first seven lines of the translated transcript of episode 1 of *Three-Body:*

♪Unity means power♪
♪Unity means power♪
♪The power is iron, the power is steel♪
♪But it is harder and stronger♪
♪Firing at Fascism, we fight♪
♪We'll wipe out all the systems
not democratic♪
♪Toward the sun, and freedom,

and new China♪[7]

The transcripts for *3 Body Problem* are also available online, through the Forever Dreaming site, which aggregates transcripts of popular TV episodes. Below are the first seven lines of the transcript of episode 1 of *3 Body Problem:*

> [crowd shouting]
> [in Mandarin] Root out the bugs!
> Sweep away all monsters and demons!
> [fervent chanting continues]
> Yes! I am a counterrevolutionary!
> [crying] I beg you
> to rehabilitate me![8]

Methods and Tools

Clearly, the transcripts for the two shows are formatted very differently. For example, the *Three-Body* transcript example starts with the lyrics of the soundtrack, denoted by the special ♪ symbol. While the opening song does set the tone for the scene, this section would likely be irrelevant to a comparison of actual dialogue, so it could be removed. The *Three-Body* transcripts also contain formatting HTML elements such as <font face="Serif" size="18">, which are similarly not relevant to the dialogue itself. Moreover, the *3 Body Problem* transcripts include sound descriptions such as [crowd shouting], which are not recorded in the *Three-Body* transcripts; thus, for a consistent comparison, they should also be removed. The aforementioned *Doctor Who* analysis makes similar choices, as "the data was cleaned of html tags and of any additional information which was not a part of the actual dialogue (comments in brackets, indications of speakers, etc.)"[9]. The transcripts also have some general quality issues – as shown

---

[7] "The Three Body Problem - 三体 (Tencent) - Google Drive," https://drive.google.com/drive/folders/1XcgTbW-GANzWvEaUkI3fZU56w5xPZti4.
[8] "3 Body Problem - Transcripts - Forever Dreaming," https://transcripts.foreverdreaming.org/viewforum.php?f=2734.
[9] Joanna Byszuk, "The Voices of Doctor Who – How Stylometry Can Be Useful in Revealing New Information About TV Series," *Digital Humanities Quarterly* 014, no. 4 (December 20, 2020).

in the *3 Body Problem* example, the sixth and seventh lines are actually from the same sentence, which has been split. In order to capture the flow of the original dialogue, these lines would need to be merged to recover the complete sentence. To handle these issues and standardize these transcripts, a tool like OpenRefine[10] would be appropriate. OpenRefine is a popular and "well-known data transformation tool"[11] that can help "clean the kinds of messy, inconsistent text data"[12] found in these transcripts to create a consistent dataset.

One method to answer the research question is a network analysis of the characters in each show. "Character networks are graphs created from the plot of a TV show that represents the interactions of characters in scenes, indicating the presence of a connection between them"[13], and they have been used in analyses of TV series to quantify character interactions – for example, a network analysis of the TV show *Friends* "looks at each character's role in the show" and provides a way of "relating each character to another"[14]. Creating character networks for *Three-Body* and *3 Body Problem* can reveal differences in how characters interact with each other across the two adaptations, as well as how the shuffling of characters in *3 Body Problem* impacts the narrative. The different group dynamics are also interesting because they may highlight how the creators of each adaptation emphasize collaboration or conflict in the story. To create these networks, a tool like Gephi[15] would be appropriate. Gephi is an open-source data

[10] "OpenRefine," https://openrefine.org/.
[11] Ruben Verborgh and Max De Wilde, *Using OpenRefine*, 2013, https://ruben.verborgh.org/publications/verborgh_packt_2013/.
[12] "Cleaning Messy Text Data Is a Breeze with OpenRefine," Eval Academy, October 15, 2020, https://www.evalacademy.com/articles/cleaning-messy-text-data-is-a-breeze-with-openrefine.
[13] Melody Yu, "Decoding the Popularity of TV Series: A Network Analysis Perspective," arXiv.org, July 4, 2023, https://arxiv.org/abs/2307.05329v2.
[14] Yusuf Sohoye, "The One with All the FRIENDS Analysis," Medium, September 11, 2019, https://towardsdatascience.com/the-one-with-all-the-friends-analysis-59dafcec19c5.
[15] "Gephi - The Open Graph Viz Platform," https://gephi.org/.

visualization tool that allows for the "easy creation of social data connectors to map community organizations and small-world networks"[16].

Another approach to answering the research question is performing topic modeling on the episodes of each show. Given a set of texts, topic modeling "discovers a set of "topics" — recurring themes that are discussed in the collection — and the degree to which each document exhibits those topics"[17]. Modeling the topics of each episode allows for a comparison of topics across time, which can help identify recurring themes and subjects throughout each series. Comparing the topics in *Three-Body* and *3 Body Problem* can uncover thematic similarities and differences, as well as reveal differences in narrative structure and plot progression. To create topic models from the show transcripts, a tool like jsLDA[18] would be appropriate. jsLDA is an in-browser tool for running Latent Dirichlet Allocation models, which are "widely applicable and approachable because it only requires a collection of raw, unannotated segments of text"[19].

Presentation and Dissemination

Ideally, the results from both the network analyses and the topic modeling approaches would be presented together on a Github page[20] that is publicly available online. The network analyses, created with Gephi, would be exported to the web using the Sigma.js exporter[21] so that visitors to the site can interact with the networks directly. The results of the topic models would be used to create line graphs spanning the timelines of the shows, where each graph corresponds

---

[16] "Gephi - The Open Graph Viz Platform."
[17] David M. Blei, "Topic Modeling and Digital Humanities," *Journal of Digital Humanities*, vol. 2, no. 1 (2012).
[18] "In-Browser Topic Modeling," https://mimno.infosci.cornell.edu/jsLDA/.
[19] Quinn Dombrowski, "Review: jsLDA," *Reviews in Digital Humanities* III, no. 8 (August 24, 2022), https://reviewsindh.pubpub.org/pub/jslda/release/1.
[20] "About GitHub Pages," GitHub Docs, https://docs.github.com/en/pages/getting-started-with-github-pages/about-github-pages.
[21] "Gephi Tutorial. Publishing Interactive Graphs Online," https://blog.miz.space/tutorial/2020/01/05/gephi-tutorial-sigma-js-plugin-publishing-interactive-graph-online/.

to a topic and there are two lines on the graph (one for each adaptation) showing how the topic has become more or less relevant in each episode. Though the two adaptations have different lengths, both generally cover the same plot points, so the lines could reasonably be scaled to the same length. This line graph format is inspired by the *Mining the Dispatch* project, which tracks topics in the Richmond *Daily Dispatch* over time[22].

In addition to the data visualizations, the site would include written summaries of the general findings of the project. The summaries would provide context and interpretations for the data visualizations and convey key insights derived from both the network analyses and topic modeling, which could potentially extend to broader cultural implications based on the narrative and thematic differences between *Three-Body* and *3 Body Problem.* Finally, to ensure transparency, the site would also include information about how the project was conducted. Copyright concerns make it difficult to provide the cleaned dataset directly, but exporting the OpenRefine project history[23] to share all the steps taken for data cleaning would allow visitors to the site to review the data cleaning process. Additionally, including information about the sources of the transcripts, the tools used in the project, and any potential limitations or biases would further increase transparency and allow visitors to the site to evaluate the project's results.

This assignment represents my own work in accordance with University regulations.

/s/ Helen Gao

[22] Robert K Nelson, *Mining the Dispatch*.
[23] "Running OpenRefine | OpenRefine," May 17, 2023, https://openrefine.org/docs/manual/running.

Bibliography

"3 Body Problem - Transcripts - Forever Dreaming." https://transcripts.foreverdreaming.org/viewforum.php?f=2734.

Blei, David M. "Topic Modeling and Digital Humanities." *Journal of Digital Humanities*, vol. 2, no. 1, 2012.

Byszuk, Joanna. "The Voices of Doctor Who – How Stylometry Can Be Useful in Revealing New Information About TV Series." *Digital Humanities Quarterly* 014, no. 4 (December 20, 2020).

Dombrowski, Quinn. "Review: jsLDA." *Reviews in Digital Humanities* III, no. 8 (August 24, 2022). https://reviewsindh.pubpub.org/pub/jslda/release/1.

Eval Academy. "Cleaning Messy Text Data Is a Breeze with OpenRefine," October 15, 2020. https://www.evalacademy.com/articles/cleaning-messy-text-data-is-a-breeze-with-openrefine.

"Gephi - The Open Graph Viz Platform." https://gephi.org/.

"Gephi Tutorial. Publishing Interactive Graphs Online." https://blog.miz.space/tutorial/2020/01/05/gephi-tutorial-sigma-js-plugin-publishing-interactive-graph-online/.

GitHub Docs. "About GitHub Pages." https://docs.github.com/en/pages/getting-started-with-github-pages/about-github-pages.

Hale, Mike. "'Three-Body' Review: A Chinese Series Beats Netflix to the Screen." *The New York Times*, February 3, 2023, sec. Arts. https://www.nytimes.com/2023/02/03/arts/television/three-body-review.html.

Heller, Jason. "'Three-Body Problem' Asks A Classic Sci-Fi Question, In Chinese." *NPR*, November 13, 2014, sec. Book Reviews. https://www.npr.org/2014/11/13/363123510/three-body-problem-asks-a-classic-sci-fi-question-in-chinese.

"In-Browser Topic Modeling." https://mimno.infosci.cornell.edu/jsLDA/.

Nelson, Robert K. *Mining the Dispatch.*

"OpenRefine." https://openrefine.org/.

Poniewozik, James. "Review: '3 Body Problem' Is a Galaxy-Brained Spectacle." *The New York Times*, March 20, 2024, sec. Arts. https://www.nytimes.com/2024/03/20/arts/television/3-body-problem-netflix-review.html.

"Running OpenRefine | OpenRefine," May 17, 2023. https://openrefine.org/docs/manual/running.

Sohoye, Yusuf. "The One with All the FRIENDS Analysis." Medium, September 11, 2019. https://towardsdatascience.com/the-one-with-all-the-friends-analysis-59dafcec19c5.

"The Three Body Problem - 三体 (Tencent) - Google Drive." https://drive.google.com/drive/folders/1XcgTbW-GANzWvEaUkI3fZU56w5xPZti4.

Verborgh, Ruben, and Max De Wilde. *Using OpenRefine*, 2013. https://ruben.verborgh.org/publications/verborgh_packt_2013/.

Yu, Melody. "Decoding the Popularity of TV Series: A Network Analysis Perspective." arXiv.org, July 4, 2023. https://arxiv.org/abs/2307.05329v2.