Subversive Characters and Stereotyping Readers: Characterizing Queer Relationalities with Dialogue-Based Relation Extraction

Kent K. Chang*, Anna Ho and David Bamman

School of Information, University of California, Berkeley, United States of America

Abstract

Television is often seen as a site for subcultural identification and subversive fantasy, including in queer cultures. How might we measure subversion, or the degree to which the depiction of social relationship between a dyad (e.g. two characters who are colleagues) deviates from its typical representation on TV? To explore this question, we introduce the task of stereotypic relationship extraction. Built on cognitive stylistics, linguistic anthropology, and dialogue relation extraction, in this paper, we attempt to model the cognitive process of stereotyping TV characters in dialogic interactions: given a dyad, we want to predict: what social relationship do the speakers exhibit through their words? Subversion is then characterized by the discrepancy between the distribution of the model's predictions and the ground truth labels. To demonstrate the usefulness of this task and gesture at a methodological intervention, we enclose four case studies to characterize the representation of queer relationalities in the *Big Bang Theory*, *Frasier*, and *Gilmore Girls* as we explore the suspicious and reparative modes of reading with our computational methods.

Keywords

conversation analysis, language models, relation extraction, television studies, gender and queer studies

1. Introduction

Television, often featuring hyper-realized characters, is an important venue for understanding social and relational identities. Take this scene from the *Big Bang Theory*:

> HOWARD. So, who wants to rent *Fiddler*?
> SHELDON. No need! We have the special edition.
> LEONARD. Well, maybe we *are* like Haroun and Tanweer. (season 1, episode 8)

Haroun and Tanweer are, as just revealed to the characters, a gay couple who recently adopted a baby. Knowing that they are a gay couple, Leonard immediately assumes that they love the musical theater that *Fiddler on the Roof* represents. Indeed, in popular culture, musical theater might have been the *queerest* genre. However, in his appraisal of its cultural significance in queer culture, John M. Clum, writing in 1999, opens with a rather curious note: "The surfeit of

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

*Corresponding author.

kentkchang@berkeley.edu (K. K. Chang); annaho@berkeley.edu (A. Ho); dbamman@berkeley.edu (D. Bamman)
0009-0008-6430-3701 (K. K. Chang)

television situation comedy has pretty much killed stage comedy. When you can get crypto-gay *Frasier* free every week, who needs the gag-rich musical comedy?" [8, p. 12]

Perhaps Clum is right. In that same year, his "crypto-gay" show aired an episode rather reminiscent of Wildean comedy of manners, in which the Crane brothers try to organize a dinner party (or according to them, an "intime soiree"), constantly on the phone inviting their friends, only to find this in the voicemail:

> ALLISON. We just got invited to a dinner party at Dr. Crane's.
>
> HARRY. Which Dr. Crane?
>
> ALLISON. Does it matter? You get the one, you get that other one. Personally, I think the whole arrangement's a little ...

What Alison truly thinks of them is never revealed to the audience, although the quick-witted Frasier jumps to his conclusion:

> NILES. What you suppose she meant by that?
>
> FRASIER. She thinks we're always together—that we're some sort of ... *couple.*
>
> NILES. That's ridiculous! We spend lots of time apart. Besides, who is she to talk? Look at her and Harry! They go everywhere together.
>
> FRASIER. They're *married*, Niles! Still, there's no reason for her to call us *odd.*
>
> (season 6, episode 17)

What ensues is another similarly frivolous argument on "who's the other one," then another on who not to invite to dinner. Much of this episode is Frasier and Niles arguing in the apartment like an old, married couple ("It's possible we have grown a tad dependent on one another.") while appreciating each other ("So we spend a lot of time together—so what? I enjoy it!").

Those aforementioned scenes from the *Big Bang Theory* and *Frasier* deal with a key aspect of discursive interactions that constitute the subject of the present study: how certain forms of dialogic interaction index certain social and relational identities, including that of a stereotypical gay man or a married couple. If Frasier and Niles are *odd* as brothers, who are they to each other? If we see their speeches as indexical signifiers, what social identities do they anchor in the interactional context? Built on cognitive stylistics [10] and social semiotics [40], this work seeks to understand how the depiction of a social relationship between a pair of characters (or a *dyad*) in conversation deviates from the typical representation of that relationship type on TV. We hope to advance an operationalization of subversion grounded in queer studies to interrogate the representation of queer relationalities on TV, and enclose four case studies to demonstrate how it can enable more close readings, which has implications for both computational humanities and queer studies.[1]

[1] See Appendix A for more related work.

2. Task and data

2.1. Task: dialogue-based stereotypic relationship extraction

Our main computational task involves predicting a stereotypical relationship type, given a dialogue between a dyad (or, two characters) from a scene in a TV series. We follow convention of relation extraction in NLP [45, 19] and refer to each dyad in terms of *head* and *tail*; unlike the core NLP task, however, we are not simply trying to optimize a model for the true relationship type, but rather identify those moments of dissonance between the truth and a prediction. For example, consider this line from *Gilmore Girls*: "Lorelai, go to your room!" Rory, who says this line, is Lorelai's daughter, but here she sounds like her *mother* for the dramatic effect, and indeed, "go to your room" sounds *stereotypically* like a parent. This kind of subversion is at the core of this work: Rory deviates from the representation of a daughter and talks like a mother. In terms of modeling, the ground truth relationship type is `child_of` for the dyad `(Rory, Lorelai)`, but we expect the prediction of the stereotypic relationship to be `parent_of`, precisely because we are modeling the stereotypic belief [10] of a viewer. Formally, given a scene $\mathcal{S}$ comprised of multiple dialogue turns, each an utterance $u_i = (s_i, w_0, \ldots, w_n)$, where $c \in S_{\mathcal{S}}$ denotes the character (or speaker) identity among all characters C in the scene $\mathcal{S}$, and w the words they speak, we seek to train a model $F(c_h^{\mathcal{S}}, c_t^{\mathcal{S}})$, where c_h is the character that occupies the head position, c_t the corresponding tail, to predict a relationship label $r \in \mathcal{R}$, making it an $|\mathcal{R}|$-way classification problem.

2.2. Dataset: dialogues and dyads

In order to carry out this inquiry, we need to identify the *true* relationship for a set of character pairs attended with dialogue in scripts, and build a model of their stereotypical interaction. To support this task, we put together a dedicated dataset of the following components:

Dialogues from parsed teleplays. We digitize and parse teleplays of pilot episodes from TV Writing to support this task.[2] Pilot episodes in the teleplay format are preferable for this task because, unlike sampling a random episode from a collection of transcripts, they typically do not require their audience to have any background knowledge of the series itself, and the standardized format of the teleplay gives us reliable scene segmentation as well as other structural information of the interaction (e.g. speaker labels and background action statements). However, digitized PDF files are unstructured. To address this, we leverage the fact that in teleplays, structural elements are distinguished by the amount of indentation a line has. For example, speaker labels are most heavily indented. We used Teseract to perform OCR on the PDF files,[3] which gives us both the recognized texts and the bounding box associated with them for each page. With this information, we used OpenAI's GPT-4o to parse those OCR'd teleplays in a one-shot setup: the model is tasked to classify each line as one of the following: scene header, speaker label, speaker note, action statement, or other, and combine consecutive lines of the same structural role. This results in 787 titles, which we split into training,

[2] https://sites.google.com/site/tvwriting/.

[3] https://github.com/tesseract-ocr/tesseract.

development, and test sets.

Relationship type labels. Another component is the relationship type labels that we can use for training models and analysis. We first use Wikipedia On-Demand API to identify and gather pages related to each title for which we have the teleplay of the pilot episode.[4] With this resource, we devise a pipeline of three stages: During the *mining* stage, we employ GPT-4o to extract relationship tuples from these summaries without using any predefined relationship labels. This extraction process was followed by a manual review, where we pruned the labels by merging semantically similar ones, such as "`kid_of`" and "`child_of`"; this is the *pruning* stage. To ensure quality, a co-author replicated the relationship extraction task on a test set comprising 50 titles in our test set. For relations that GPT-4o did extract, the accuracy is satisfactory at 81.67%. Finally, we verify the accuracy of all extractions in both the development and test sets, which concludes the *verification* stage. In principle, several of the relationship types can be overlapping—two characters may be seen to be both married (`spouse_of`) and lovers (`love_interest_of`). In order to create a multiclass classification problem (predicting only one label for each dyad), we rank all relationship types based on their specificity (as listed in Table 3 in Appendix B) and ask a model to predict the most *specific* relationship type from this set. The most frequent relationship types are: `colleague_of`, `friend_of`, `sibling_of`, `spouse_of`, and `classmate_of`.

Dyads. To generate dyads for training and analysis, we begin by collecting the set of speaker labels present in each teleplay. These labels are often noisy due to OCR errors and inconsistencies with how names appear in Wikipedia summaries. To standardize these labels, we query the title in the Movie Database (TMDb),[5] which provides a list of characters, both recurring and guest. We then match each speaker label with a canonical character name from TMDb using a simple heuristic: if there is at least one token overlap, we select the TMDb character name with minimal edit distance. Labels that do not match (typically generic names like "MAN #1") are discarded. With standardized speaker labels, we create the list of dyads of interest for each scene by permuting the labels in pairs (i.e. creating 2-permutations from the set of distinct speakers in a scene). This means that a pair of characters is considered only if both have speaking roles within the same scene. Finally, we include a dyad in our dataset if we have previously extracted its relationship type from Wikipedia. See Appendix C for statistics and more information.

Anonymization. Previous work indicates speaker anonymization is beneficial as it mitigates the distraction and noise named entities might introduce [7]. We similarly anonymize our dataset to evaluate its impact. For each scene, we maintain a mapping table between canonical speaker names and randomly assigned entity IDs. All speaker identities are then replaced with `ENTITY` plus their unique scene ID, and their mentions in dialogue lines are anonymized in the same way. Some canonical names include generic words, often related to a profession (e.g. "coach"), and those words would be anonymized, which might lead to an unnecessary

[4] https://enterprise.wikimedia.com/products/.

[5] https://www.themoviedb.org/.

```
<background> Briefcase in hand , ENTITY 5 is once again waiting for the elevator . He 's
    approached by ENTITY 6 , 39 , smart , cute , but not sweet . You do n 't get to be
    hospital administrator and dean of medicine by being sweet .
<speaker> ENTITY 6 <line> I was expecting you in my office twenty minutes ago .
<speaker> ENTITY 5 <line> Really ? That 's odd because I had no intention of being in
    your office twenty minutes ago .
```

Figure 1: Example of an anonymized and post-processed scene (only first two lines represented here).

loss of information. To address this, we aggregate all tokens in canonical speaker names and manually annotate whether each token is part of a proper name or a non-name content word. An example is in Fig. 1 (special tokens like `<speaker>` are explained in Sec. 3).

3. Models and experiments

Given this data, we can build and evaluate models of *stereotyping readers*, who learn from the totality of relationships encoded in our training data to infer the relationship enacted between two characters in a specific scene. We do not expect a model to be able to identify the *true* relationship with perfect accuracy—not every relationship is enacted in dialogue at the level of a single scene; but more importantly, our work is premised on the idea that the relationship we observe on screen (and that a model observes as well) can exhibit variation that is deliberately at odds with that "truth." And yet, accuracy at predicting that truth provides us with an instrumental means to select between different models, and we assess several variants in order to find the one most sensitive to the dialogic indicators of how relationships are performed.

3.1. Models

We compare the performance of the following models on this task to select the best strategy to model stereotyping readers with our data. We first establish the **majority class** (predicting the most frequent class, `colleague_of`) baseline for all models. Since the task takes the form $F(c_h^{\mathcal{S}}, c_t^{\mathcal{S}})$, those models can be categorized based on how we choose to represent the character c_h and c_t in scene $\mathcal{S}$: *supervised* and *prompting*:

Supervised. We start with an *utterance-based* model, where we string together all utterances spoken by the head and tail speakers, respectively. The motivation is that each speaker can be represented by all of their utterances in the scene. Those utterances are subsequently encoded by the same Longformer–base encoder [2]; we extract the CLS tokens (`<s>`) that represent all the utterances and concatenate them, resulting in the overall representation:

$$\boldsymbol{h} = [\boldsymbol{e}^{c_h}_{\texttt{<s>}}; \boldsymbol{e}^{c_t}_{\texttt{<s>}}]. \tag{1}$$

Next, we include a representation of the entire scene using the *attentive pooling* technique. Here, we similarly string all the tokens in the entire scene together, but to enhance the structural awareness of the teleplay (e.g. some tokens are speaker labels, and some are their lines),

Table 1
Experimental results. All metrics are reported with 95% bootstrap confidence intervals.

	Accuracy	
	anonymized test set	unanonymized test set
Majority	0.265	0.265
SUPERVISED		
Longformer	0.305 [0.294–0.315]	0.298 [0.287–0.309]
+ anonymized training set	0.293 [0.283–0.303]	0.306 [0.295–0.317]
+ scene attentive pooling	0.248 [0.238–0.258]	0.348 [0.337–0.359]
+ both	**0.338** [0.327–0.349]	**0.367** [0.356–0.378]
PROMPTING		
LLaMA 3–70b	0.197 [0.188–0.206]	0.243 [0.233–0.253]
+ one-shot	0.212 [0.202–0.221]	0.242 [0.232–0.252]
OpenAI o1-mini	0.181 [0.172–0.190]	0.241 [0.231–0251]

we introduce the following special tokens, whose representations are learned during training: `<scene>`, `<speaker>`, `<line>`, and `<background>`. A full example is in Fig. 1. We take inspiration from [35] and incorporate attentive pooling techniques for the scene representation. Since we have the utterances from both head and tail speakers from the utterance-based model, we want to emphasize other information in the scene by guiding the model to attend less to those utterances and more to dialogue lines from other speakers. In encoding the scene here, we introduce a token-level mask M, where $M[j] = 0$ if the j-th word is spoken by either head or tail speaker and $M[j] = 1$ otherwise. Following [35], the scene information selected by M is:

$$\boldsymbol{h}_{\mathcal{S}} = \boldsymbol{e}^{\mathcal{S}}_{\texttt{<s>}}; \qquad A = \boldsymbol{w}_A{}^{\top}\boldsymbol{h}_{\mathcal{S}}; \qquad \alpha = \mathrm{softmax}(A \odot M). \tag{2}$$

The head- and tail-aware attention is used to pull the hidden states: $\boldsymbol{h}_{\mathcal{S}}{}^{\top}\alpha$, which is concatenated with head and tail utterances:

$$\boldsymbol{h} = [\boldsymbol{e}^{c_h}_{\texttt{<s>}}; \boldsymbol{e}^{c_t}_{\texttt{<s>}}; \boldsymbol{h}_{\mathcal{S}}{}^{\top}\alpha]. \tag{3}$$

The overall representation $\boldsymbol{h}$ is then fed to a linear classification head f, which yields $P(r|\boldsymbol{h}) = \mathrm{softmax}\big(f(\boldsymbol{h})\big)$ and $\hat{r} = \arg\max P(\cdot)$.

Prompting. For prompt-based models, the overall prompt design resembles a QA task, where given a scene, the model is tasked to answer, *head speaker is ____ of tail speaker.* We consider three popular strategies to enhance the performance of large language models: we prompt LLaMA 3–70b–instruct zero-shot and one-shot,[6] and leverage the hidden chain-of-thought process [44] in OpenAI's o1-mini.[7] For more details, see Appendix D.

Table 2
Most distinct words measured by log-odds ratio for key relationship types.

Relation type	Most distinct words
`parent_of`	kids, son, mother, house, father, darling, honey, debate, worried, stay
`sibling_of`	sister, brother, whistledown, lady, cherry, lord, mom, dollars, hastings, must
`spouse_of`	honey, love, kids, marriage, baby, married, clean, treat, maple, care
`colleague_of`	death, find, magic, heroin, found, real, ship, missing, library, case
`friend_of`	girls, fun, york, president, buddy, school, high, rally, jacket, vote

3.2. Experimental results

Experimental results are reported in Table 1, including 95% confidence intervals from 10,000 bootstrap resamples.[8] We evaluate the performance of each model by comparing its prediction against the true relationship label we have obtained from Wikipedia. Given the premise of this work is that the true relationship type is not necessarily performed in every interaction, we expect the accuracy here to be low, but the model should still perform better than guessing the most frequent label in the training set. In assessing the impact of anonymization, we include additional rows for when we anonymize the training set and columns for test ones.

For the supervised models, we observe that adding contextual information about the scene that is absent from the head and tail utterances significantly improves the performance of the model. When the model already has the scene information, anonymizing the training set appears beneficial, which is aligned with the findings reported in [7]. For evaluation, anonymization does not significantly hurt the performance of our supervised models. However, it does for the prompt-based models we evaluate. While they yield similar performance on the unanonymized data (around 0.242), the three prompt-based models do not produce identical predictions: LLaMA's zero-shot and one-shot models have a Cohen's κ [9] of 0.71[0.70, 0.72], and that between LLaMA zero-shot and OpenAI o1-mini is 0.37[0.36, 0.39]. This suggests a low agreement rate between LLaMA and OpenAI models.

Since the goal of this section is to figure out the best strategy for modeling stereotyping readers, it is crucial to establish the facial validity of the best-performing model before we impart any trust in its predictions. In further examining the face validity of the predictions of our best-performing model before moving on to the analysis, we represent the most distinct tokens for the key *predicted* relationship types measured by log-odds ratio with an uninformative Dirichlet prior [30] in Table 2, which we can consider a form of post-hoc global explanation providing insight into what a model has learned [12]. We aggregate all head utterances by the predicted relationship type and use them as the target corpus, and the rest as the reference corpus. For each relationship type presented, we include the top ten tokens most strongly asso-

[6] https://huggingface.co/meta-llama/Meta-Llama-3-70B.

[7] https://openai.com/index/openai-o1-mini-advancing-cost-efficient-reasoning/.

[8] For supervised models, we use a learning rate of 5×10^{-5} with 100 warm-up steps and no weight decay. All inputs are padded or truncated to 4, 096 tokens. All models are trained on four L40S GPUs. If the scene is longer than 4, 096 tokens, it is truncated before being inserted into the prompt (only four scenes were truncated). We use Outline to constrain the output space to be one of the relationship types: https://github.com/outlines-dev/outlines.

ciated with the target corpus. Those five types are chosen because they are central to our case studies presented in the analysis section below. In Table 2, we see that the model has learned to associate the relationship types with some of the words the dyad in question typically talks about (e.g. parents talk about kids), *colleague* has to do with occupational terms on TV (e.g. detectives investigate the death of someone), and *friend* is focused on high school life (e.g. students vote for student council president). Although there are terms that seem confusing (say, *jacket* in `colleague_of`) out of context, the words that are more strongly associated with those categories make an intuitive sense in the context of scripted TV series.

4. Analysis

> Well, I suppose I do think of you as a sister. And sometimes, a mother.
> —Sheldon Cooper to his friend, Penny, in the *Big Bang Theory*

> [O]ne of the things that "queer" can refer to: the open mesh of possibilities, gaps, overlaps, dissonances and resonances, lapses and excesses of meaning [. . .].
> —EVE KOSOFSKY SEDGWICK, *Tendencies* [36]

Television is often seen as a medium for subcultural identification [16] and subversive fantasy [31], and this work is animated by the sociological impulse in certain strands of queer studies that see artistic forms as reflecting "the texture and makeup of queer social worlds" [26, p. 115]. From this perspective, we can see certain moments on TV as representing a queer time and place that sits "in opposition to the social institutions of family, heterosexuality, and reproduction" [18, p. 1], where endless possibilities of subversive relational forms take shape in the "excesses of meaning" [36, p. 8]. We can use the model described in this paper to characterize this phenomenon. For our analysis, we choose the dataset introduced in Sang, Mou, Yu, Yao, Li, and Stanton [35], which consists of five TV series, each almost in their entirety, which allows us to study narrative arcs that span multiple seasons. Through the case studies, we hope to demonstrate how this work might shed light on the representation of queer modes of relating in the *Big Bang Theory*, *Frasier*, and *Gilmore Girls.*

4.1. Queer characters and suspicious reading

Our first case study takes up David Halperin's inquiry into *queer love* [17] and its central question: "how is it possible for two men to be together" when existing social institutions cannot accommodate such a form of togetherness [14, p. 136]. For Michel Foucault and Halperin, the love between men is queer not because of their sexual preferences and practices; it is instead because of their counter-conduct [11]. In the context of this study, for any given discursive and dyadic interaction on television, we see how "one conducts oneself, lets oneself be conducted, and finally, in which one behaves under the influence of a conduct as the action of conducting" [11, p. 128]. Conduct is dictated by the social relationship the dyad indexically presumes and projects through conducting with their speech acts. In this light, the act of subversion through speech—when characters disrupt expected linguistic norms by using forms typically reserved for certain social categories—transgresses the presumptive bounds of relational form and is essential to the practice of counter-conduct. Importantly, counter-conduct necessitates

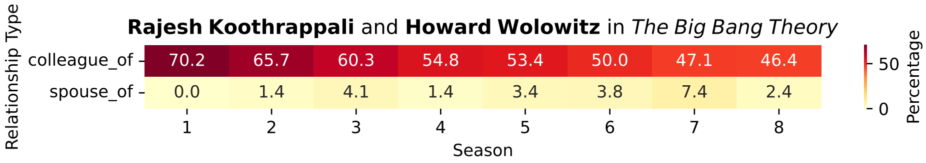


Figure 2: A heatmap representing the progression of relationship types, "`colleague_of`" and "`spouse_of`", between Raj and Howard in *The Big Bang Theory* across the first eight seasons.

new forms of relationality: men in queer love need to "invent, from A to Z, a relationship that is still formless" [14, p. 136]. Can we use our model to investigate counter-conduct and relational forms? The titles in our analysis dataset feature some characters that have become the subject of various queer readings: for example, Frasier and Niles from *Frasier* are said to embody gay sensibility [33], and the four male protagonists of the *Big Bang Theory* perform "queer-straight masculinity" [29]. Juxtaposing the relational possibilities queerness affords and the subversive potentials that many find inherent in those characters, we can use our model as a tool to facilitate and augment close reading.

In the *Big Bang Theory*, Raj Koothrappali and Howard Wolowitz are colleagues at Caltech, but they often talk like a couple:

> HOWARD. We're just saying all the things we love about each other.
> RAJ. Oh, like you and I did at couple's therapy? (season 8, episode 9)

Much debate surrounding Raj focuses on whether he is gay, to which Steve Molaro, producer and writer of the show, says: while it's viable, "it was a little more interesting to have a guy so comfortable in his feminine side who's not gay, and explore that" [32, p. 244]. Along with his self-identification as a metrosexual and fascination with divas ("Cher, Madonna, Adele. All the women who rock me", season 5 episode 14), we might understand Molaro as gesturing at separating two distinct dimensions of the Raj character: his sexual orientation and, per Foucault, his "way of life" [14]. Against this backdrop, we argue that Raj and Howard offer more than occasional punchlines; they represent an instance of queer love: they have to invent a form of togetherness for themselves, despite being constantly under the watch of, occasionally derided by, the maternal signifier, the figure sans figure Mrs. Wolowitz, Howard's mother, who we never see on screen: "Frankly, after all your sleepovers with the little brown boy, a girl is a big relief" (season 5, episode 3). This need to *invent* is made explicit by another character in the series, Dr. Beverly Hofstadter, albeit in a pejorative tone: "the two of you have created an ersatz homosexual marriage to satisfy your need for intimacy" (season 2, episode 15).

In Fig. 2, we chart the percentage of the dyad being predicted as *colleagues* or *spouses* to each other across the eight seasons in our dataset of choice. It is not surprising to see *colleague* being the most prominent relationship type, but it gradually decreases from 70.2% in season 1 to 46.4% in Season 8, while the "spouse_of" relationship type is predicted starting from season 2, peaking at 7.4% in Season 7. This indicates a predominant colleague relationship with minor fluctuations towards a more intimate or spousal-like dynamic. While talking like spouses when they are colleagues can be seen as an act of transgression, do those characters

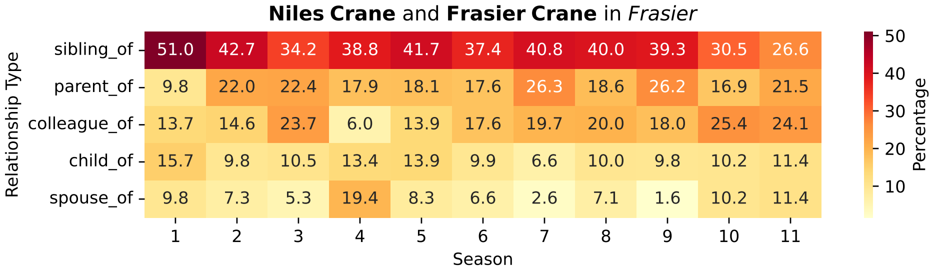


Figure 3: Another heatmap representing the progression of relationship types between Niles and Frasier in *Frasier* over eleven seasons. Only the top five relationship types are presented here.

necessarily submit to the social institution of marriage in their form of co-existence, which the *spouse* category would entail? If we believe our stereotyping reader and assume that the spouse is an observable relational form between them, we ask: can there be subversion within subversion, and do Raj and Howard contest this normative relational form?

For that we return to close, suspcious reading. One relevant episode takes place towards the end the series (season 12, episode 22): We find Raj at an airport, waiting for his flight to London, where he is expected to meet with his girlfriend, Anu. And we see Howard in discussion on this with his wife, Bernadette:

> BERNADETTE. Go stop him. Get your best friend back.
> HOWARD. You *are* my best friend!
> BERNADETTE. We don't have time for this! Go!

This exchange, along with the ensuing airport scene, stages a jubilant celebration of queer love. If, according to Halperin, "where the happy couple advances, deviance retreats" [17, p. 397], thanks to Bernadette, we see this playing out in the opposite direction: the married couple steps aside, for queer love to flourish. Howard getting Raj back by no means signals the end of his marriage with Bernadette, but it surfaces the tension between traditional heteronormative relationships and non-normative desires: Howard's dynamic with Raj reveals an undercurrent of emotional intimacy and dependency that challenges the rigid boundaries of what is considered acceptable, *normal* male bonding within the confines of marriage. Bernadette, curiously, becomes a kind of referee, simultaneously reinforcing and undermining the normalcy of her marriage: She exposes the instability of the heteronormative relationship model, while also ensuring that it doesn't just break apart. As Oscar Wilde puts it in his putatively queer *Earnest*, "In married life, three is company and two is none."

Now we return to Niles and Frasier Crane, to the question we raise at the beginning of this paper: If they are not *just* brothers, who are they to each other? It is not surprising that in Fig. 3, the most salient relationship type for them is that of a sibling, but we see how they perform a few different ones over the course of the show. Towards the end, Nile appears to be just as much a parent and a colleague as a brother for Frasier. We see how they "invent" (in Foucault's word) a new relational form for themselves as they oscillate between the relationship types we

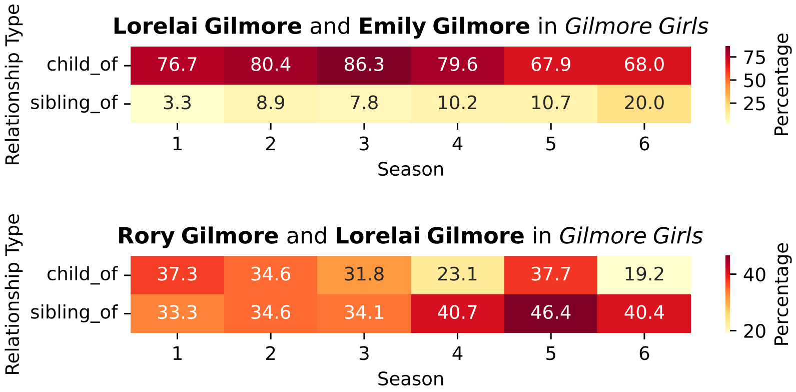


Figure 4: Heatmap representing the mother–daughter relationship arc between Rory and Lorelai Gilmore (above) and between Lorelai and Emily Gilmore (below) in *Gilmore Girls* in the first six seasons.

already have names for, although others, like Alison and Harry from Introduction, might find them odd—and, indeed, queer—at times.

4.2. Anti-normative characters and reparative reading

Today, *queer* encompasses a broad spectrum of identities, practices, and desires related to sex and gender. But this is not the case for queer theory at its infancy: writing around 1993, Eve Kosofsky Sedgwick notes how queer scholarship "can't be subsumed under gender and sexuality at all" [36, p. 8]. In this section, we operate on this expanded notion of queer and turn to subversive modes of parenting.

Recall this line from Section 2: "Go to your room, Lorelai" (season 1, episode 4) is the daughter talking like a mother. This is an instance of subversion where traditional parent–child boundaries is blurred, as Rory no longer conforms to the conventional roles. We do see two modes of parenting depicted in *Gilmore Girls*: the traditional model Lorelai and her mother, Emily Gilmore represents, and the new, subversive one between Rory and Lorelai. In Fig. 4, we chart the interaction patterns by the percentage of the predicted relationship type between these two pairs of mother and daughter each season, focused on the top three relationship types. According to our model, we see Emily, compared to Lorelai, is more of a traditional mother throughout, with `parent_of` being the dominant relationship type. This is aligned with our impression with the characters: as Lorelai says, "Rory and I are best friends, Mom. We're best friends first, and mother and daughter second. And you and I are mother and daughter always" (season 2, episode 16). However, as we see in Fig. 4, they start to talk more like sisters around each other in later seasons. One episode where the dynamic between Lorelai and Emily changes, as identified by our model is "Friday and Alright For Fighting" (season 6, episode 13), where Emily says jokingly to Lorelai, "I only wished I'd remembered to call her a cocktail waitress!" This is a surprise to Lorelai: "That's my mother's version of the *c* word!"

Another intriguing example of subversive parenting can be found between Sheldon Cooper and Penny in the *Big Bang Theory*. In this show, Sheldon works with Raj and Howard at Caltech,

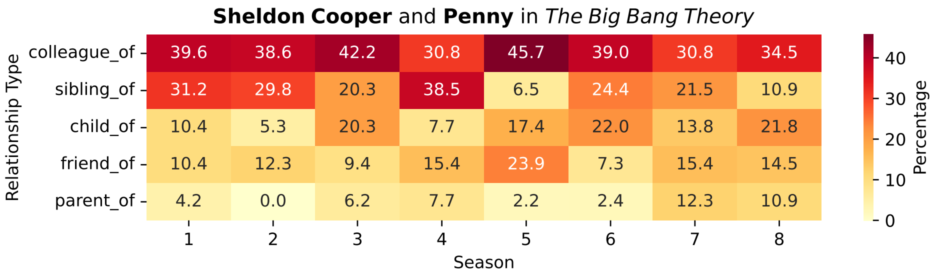


Figure 5: Another heatmap representing the progression of relationship types between Sheldon Cooper and Penny over eight seasons in the *Big Bang Theory*. Only the top five relationship types are presented here.

and Penny represents the "cute girl next door next to the nerds" [13] archetype: the main male cast routinely succumbs to her feminine charm, especially in early seasons, with the sole exception of Sheldon. For this reason, Sheldon is regarded as asexual [29] and infantilized [38]. The latter dimension of the character is explored throughout the series through Penny, which is the main plot of "The Intimacy Acceleration" (season 8, episode 15), from which the epigraph of this section is taken. In this episode, the two engage in a farcical experiment to find out if Sheldon and Penny can fall in love. For Sheldon, his announced goal is not to win the girl, but to have her drive him to a "convention celebrating the life and work of Gary Gygax." Their experiment commences with Penny saying, "I will buy you all the dragon T-shirts you want." For most of this episode, like a Kleinian baby [20], Sheldon explores object relations and symbol formation in his phantasy-as-experiment: at the end of it, he compares himself to "human bowl of tomato soup", and Penny to his sister and his mother.

Our model of stereotyping readers captures that, and more. The top relationship types for Sheldon and Penny do include `sibling_of` and `child_of`, and according to the model, one of the moments where Sheldon talks like Penny's child involves her trying to put him to sleep, though unsuccessfully: "No, I don't want to go to sleep, you can't make me" (season 8, episode 13). What's surprising here is Sheldon and Penny, much like Frasier and Niles, take up multiple relational roles over the eight seasons, and the most frequent of them is that of *colleagues*. If we revisit the series with a keen eye on when they speak like colleagues, we see how Penny can act like a counsel in Sheldon's social and everyday life: on how to empathize with others (season 2, episode 3), girl trouble at a bar where Penny works (season 4, episode 17), and so on. Penny's capacity for social mentoring reverses the normal dynamics between them, where Sheldon feels and acts like her superior. In terms of intertextuality, this can explain why Sheldon thinks of Penny as a sister: in the prequel to the series, *Young Sheldon*, we see his sister, Missy Cooper, possess extraordinary social intelligence.

Surprise is the operative word for this section, and this choice of words is intentional. Inspired by Melanie Klein's work, including her theories on object relations and phantasy, Sedgwick developed her concept of repatative reading: "[T]o read from a reparative position is to surrender the knowing, anxious paranoid determination that no horror, however apparently unthinkable, shall ever come to the reader as *new*; to a reparatively positioned reader, it can seem realistic and necessary to experience surprise" [37, p. 146]. In being surprised, we see how computational methods can enable us to take the reparative position as we approach the text; we attempt to bring out the multiplicity of characters, like Emily and Penny, and begin to repair their agency. Indeed, they are much more than stereotypical mothers or stereotypical cute girls.

5. Conclusion

> A conversation has no necessary terminus.
> —TYLER BRADWAY, "Queer Narrative Theory and the Relationality of Form" [3]

In this paper, we model a stereotyping reader who can infer the social relationship of pairs of characters (or dyads) in conversation. As an algorithmic measuring device, this model is queer in and of itself: the model tries to learn, for example, what a couple talks like, but we are ultimately interested in finding characters that are *not* a couple but talk like one, therein lies the queerness we wish to explore. As such, metrics like accuracy (as we argue in Section 3) are at best an instrumental means, and the model is interesting (as we see in the case studies) only when it predicts anything *but* the truth. As queer studies intersect with computational humanities, through our model, we wish to take initial steps to "dismantle the logics of success and failure" [15, p. 2] to resist the "normal business in the academy" [43, p. xxvi], and in so doing, reflect on disciplinary practices in the relevant research communities to "build a better description" [28] of queer culture.

One of the powerful images that Sedgwick invokes in her work, which has ultimately altered the landscape of queer theory permeably and forever, is that of a "theory kindergarten" [37, p. 94].[9] As we gesture towards a potential future for a queer cultural analytics [6], we look back on Sedgwick's theory kindergarten. Sedgwick was critical, but we might think of a theory kindergarten as a generative space of playful exploration, not bound by rigid epistemological frameworks or normative, institutional constraints. As such, theory kindergarten can be as inclusive, reflexive, and enabling: as we see in *Touching Feeling*, it invites us to "think otherwise" [37, p. 11]: embrace surprise and creativity, allow for new modes of understanding and inquiry to emerge—perhaps including computation, without reducing the complexity and multiplicity of queer thought. If relational forms that emerge out of dialogic interactions showcase the "open mesh of possibilities" of queerness [36, p. 7] that Sedgwick speaks of, interdisciplinary conversations, as Bradway, the source of the epigraph for this section, so eloquently intimates in the context of queer narrative theory, do not need an overdetermined terminus. We hope this work can motivate more researchers to study the representation of subversion in queer cultures and test the limits of both queer theory and computational methods. Back in the kindergarten: experiment and play, think and operationalize otherwise, and reimagine queerness vis-à-vis computation.[10]

[9] See also [4], [25].

[10] Code to support this work can be found at: https://github.com/kentchang/subversive-characters-chr2024.

Acknowledgments

We thank the anonymous reviewers for their helpful comments. We are grateful for feedback provided by Richard So and Mackenzie Cramer at different stages of this project. The research reported in this article was supported by funding from the National Science Foundation (IIS-1942591), the National Endowment for the Humanities (HAA-271654-20), and a UC Berkeley Paul Fasana LGBTQ Studies Fellowship to K. C.

References

[1] M. Bednarek. *Language and Characterisation in Television Series: A Corpus-informed Approach to the Construction of Social Identity in the Media.* John Benjamins Publishing Company, 2023.

[2] I. Beltagy, M. E. Peters, and A. Cohan. "Longformer: The Long-Document Transformer". In: (2020). arXiv: 2004.05150 [cs.CL].

[3] T. Bradway. "Queer Narrative Theory and the Relationality of Form". In: *Publications of the Modern Language Association of America* 136.5 (2021), pp. 711–727.

[4] D. P. Britzman. "Theory Kindergarten". In: *Regarding Sedgwick.* Ed. by S. M. Barber and D. L. Clark. London, England: Routledge, 2013, pp. 121–142.

[5] J. Butler. *Gender Trouble: Feminism and the Subversion of Identity.* Ny: Routledge, 1990.

[6] K. K. Chang. "The Queer Gap in Cultural Analytics". In: *Debates in the Digital Humanities 2023.* Ed. by M. K. G. Lauren F. Klein. U of Minnesota Press, 2023, pp. 105–119.

[7] M. Chen, Z. Chu, S. Wiseman, and K. Gimpel. "SummScreen: A Dataset for Abstractive Screenplay Summarization". In: *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers).* Dublin, Ireland: Association for Computational Linguistics, 2022, pp. 8602–8615.

[8] J. M. Clum. *Something for the Boys: Musical Theater and Gay Culture.* St. Martin's Press, 2001.

[9] J. Cohen. "A Coefficient of Agreement for Nominal Scales". In: *Educational and psychological measurement* 20.1 (1960), pp. 37–46.

[10] J. Culpeper. *Language and Characterisation: People in Plays and Other Texts.* Longman, 2001.

[11] A. I. Davidson. "In Praise of Counter-Conduct". In: *History of the human sciences* 24.4 (2011), pp. 25–41.

[12] M. Du, N. Liu, and X. Hu. "Techniques for Interpretable Machine Learning". In: *Communications of the ACM* 63.1 (2019), pp. 68–77.

[13] A. Dumaraog. *Big Bang Theory: Kaley Cuoco Explains How Penny Became Less Sexualized.* https://screenrant.com/big-bang-theory-penny-sexualized-kaley-cuoco-response/. 2021.

[14] M. Foucault. “Friendship as a Way of Life”. In: *Ethics: Subjectivity and Truth (Essential Works of Foucault, 1954–1984, Vol. 1).* Ed. by P. Rainbow. NY: The New Press, 1998, pp. 135–140.

[15] J. Halberstam. *The Queer Art of Failure.* Duke University Press, 2011.

[16] D. M. Halperin. *How To Be Gay.* Harvard University Press, 2012.

[17] D. M. Halperin. “Queer Love”. In: *Critical inquiry* 45.2 (2019), pp. 396–419.

[18] J. Jack Halberstam and J. Halberstam. *In a Queer Time and Place: Transgender Bodies, Subcultural Lives.* NYU Press, 2005.

[19] Y. Jiang, Y. Xu, Y. Zhan, W. He, Y. Wang, Z. Xi, M. Wang, X. Li, Y. Li, and Y. Yu. “The CRECIL Corpus: A New Dataset for Extraction of Relations between Characters in Chinese Multi-Party Dialogues”. In: *Proceedings of the Thirteenth Language Resources and Evaluation Conference.* Ed. by N. Calzolari, F. Béchet, P. Blache, K. Choukri, C. Cieri, T. Declerck, S. Goggi, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, J. Odijk, and S. Piperidis. Marseille, France: European Language Resources Association, 2022, pp. 2337–2344.

[20] M. Klein. *Love, Guilt, and Reparation & Other Works, 1921–1945.* Her the Writings of Melanie Klein. New York, NY: Delacorte Press, 1975.

[21] S. Kozloff. *Overhearing Film Dialogue.* University of California Press, 2000.

[22] G. Li, Z. Xu, Z. Shang, J. Liu, K. Ji, and Y. Guo. “Empirical Analysis of Dialogue Relation Extraction with Large Language Models”. In: (2024). arXiv: 2404.17802 [cs.CL].

[23] X. Liu, J. Zhang, H. Zhang, F. Xue, and Y. You. “Hierarchical Dialogue Understanding with Special Tokens and Turn-Level Attention”. In: *Tiny Papers ICLR* (2023).

[24] H. Love. “Doing Being Deviant: Deviance Studies, Description, and the Queer Ordinary”. In: *Differences* 26.1 (2015), pp. 74–95.

[25] H. Love. “Truth and Consequences: On Paranoid Reading and Reparative Reading”. In: *Criticism* 52.2 (2010), pp. 235–241.

[26] H. Love. *Underdogs.* Chicago, IL: University of Chicago Press, 2021.

[27] B.-R. Lu, Y. Hu, H. Cheng, N. A. Smith, and M. Ostendorf. “Unsupervised Learning of Hierarchical Conversation Structure”. In: *Findings of the Association for Computational Linguistics: EMNLP 2022.* Ed. by Y. Goldberg, Z. Kozareva, and Y. Zhang. Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, 2022, pp. 5657–5670.

[28] S. Marcus, H. Love, and S. Best. “Building a Better Description”. In: *Representations* 135.1 (2016), pp. 1–21.

[29] A. McClanahan. “Disciplining Heterosexuality: Interrogating the Heterosexual Ideal”. In: *The Sexy Science of The Big Bang Theory: Essays on Gender in the Series.* Ed. by N. Farghaly and E. Leone. McFarland, 2015, pp. 88–110.

[30] B. L. Monroe, M. P. Colaresi, and K. M. Quinn. “Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict”. In: *Political analysis: an annual publication of the Methodology Section of the American Political Science Association* 16.4 (2017), pp. 372–403.

[31] T. Pugh. *The Queer Fantasies of the American Family Sitcom*. Rutgers University Press, 2018.

[32] J. Radloff. *The Big Bang Theory: The Definitive, Inside Story of the Epic Hit Series*. London, England: Grand Central Publishing, 2022.

[33] D. Raymond. "Popular Culture and Queer Representation". In: *A Critical Perspective* (2003), pp. 98–110.

[34] K. Richardson. *Television Dramatic Dialogue: A Sociolinguistic Study*. Oxford University Press, 2010.

[35] Y. Sang, X. Mou, M. Yu, S. Yao, J. Li, and J. Stanton. "TVShowGuess: Character Comprehension in Stories as Speaker Guessing". In: *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Ed. by M. Carpuat, M.-C. de Marneffe, and I. V. Meza Ruiz. Seattle, United States: Association for Computational Linguistics, 2022, pp. 4267–4287.

[36] E. K. Sedgwick. *Tendencies*. Duke University Press, 1993.

[37] E. K. Sedgwick. *Touching Feeling*. Duke University Press, 2003.

[38] J. Shaw. "The Adolescent Quest". In: *The Sexy Science of The Big Bang Theory: Essays on Gender in the Series*. Ed. by N. Farghaly and E. Leone. McFarland, 2015, pp. 72–87.

[39] M. Silverstein. ""Cultural" Concepts and the Language-Culture Nexus". In: *Current anthropology* 45.5 (2004), pp. 621–652.

[40] M. Silverstein. *Language in Culture: Lectures on the Social Semiotics of Language*. Cambridge University Press, 2022.

[41] L. C. Stache and R. D. Davidson. *Gilmore Girls: A Cultural History*. Rowman & Littlefield, 2019.

[42] Q. Sun, B. Schiele, and M. Fritz. "A Domain Based Approach to Social Relation Recognition". In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI: Ieee, 2017.

[43] M. Warner. *Fear of a Queer Planet*. Studies in Classical Philology. Minneapolis, MN: University of Minnesota Press, 1993.

[44] J. Wei, X. Wang, D. Schuurmans, M. Bosma, B. Ichter, F. Xia, E. Chi, Q. V. Le, and D. Zhou. "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models". In: *Advances in Neural Information Processing Systems*. Ed. by S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh. Vol. 35. Curran Associates, Inc., 2022, pp. 24824–24837.

[45] D. Yu, K. Sun, C. Cardie, and D. Yu. "Dialogue-Based Relation Extraction". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Ed. by D. Jurafsky, J. Chai, N. Schluter, and J. Tetreault. Online: Association for Computational Linguistics, 2020, pp. 4927–4940.

A. Additional related work

This work builds on prior research in several disciplinary traditions.

Sociolinguistics and linguistic anthropology. The study of social relationships through language is central to the fields of sociolinguistics and linguistic anthropology. These disciplines provide tools for analyzing how language constructs and reflects social identities and power dynamics. For this work, Michael Silverstein's work [39, 40], which explores the ways in which language ideologies and linguistic practices intersect, informs our approach to how social relationships are subverted and maintained through dialogue.

Dialogue understanding and natural language processing. The design and implementation of our computational methods are indebted to the field of natural language processing (NLP), especially work on dialogue understanding. Our main task is built on dialogue relation extraction techniques, which are employed to classify the relationships between characters based on their conversational exchanges [19, 22]. This work also builds upon narrative understanding and conversation modeling. This is exemplified by, among others, TVShowGuess [35], which leverages neural models to perform reading comprehension tasks on television shows. In the context of this work, understanding the hierarchical structure of dialogue is also crucial [42, 27, 23].

TV and film studies. Much of this work aims to understand representation on screen, which is the subject of TV and film studies [41], and in particular, analyzing dialogue within TV and film is essential for understanding how social relationships are portrayed and subverted. This work is inspired by linguistic analysis in this space [34, 1] and works such as Sarah Kozloff's [21] which look into the mechanics of scripted dialogue and its impact on the audience's perception of characters and their relationships.

Gender and queer studies. The subversion of social relationships in TV often intersects with issues of gender and sexuality. Judith Butler's theories on gender performativity provide the essential framework for understanding how characters on TV subvert traditional gender roles and expectations [5]. The positivist approaches in queer studies [16, 24] inform the theoretical foundation of this work as we articulate subversion in computational terms.

B. Ranked relationship types

See Table 3.

C. Dataset statistics

After the process described in Section 2, each pilot teleplay is now transformed into a sequence of scenes, which is comprised of speakers with canonical names and their lines, as well as,

Table 3
Ranked relationship types.

Rank	Relationship Type	Rank	Relationship Type
1	grandparent_of	15	enemy_of
2	grandchild_of	16	colleague_of
3	parent_of	17	classmate_of
4	child-in-law_of	18	roommate_of
5	child_of	19	neighbor_of
6	sibling-in-law_of	20	teacher_of
7	sibling_of	21	student_of
8	relative_of	22	boss_of
9	ex-spouse_of	23	subordinate_of
10	ex-boy/girlfriend_of	24	trainer_of
11	ex-love_interest_of	25	trainee_of
12	spouse_of	26	acquaintance_of
13	boy/girlfriend_of	27	friend_of
14	love_interest_of	28	other

where applicable, a list of relationship tuples for speakers in the scene. The statistics of title and token counts for this dataset are reported in Table 4.[11]

Table 4
Summary of train, development, and test sets.

	training	development	test
# titles	552	115	120
# scenes	36,320	7,078	6,965
# number of dyads with labels	9,223	4,978	7,176
# avg tokens per			
scene	192	194	201
utterance	95	96	103

D. Sample prompts

D.1. LLaMA 3 prompt

See Fig. 6.

D.2. Sample OpenAI o1-mini prompt and response

See Fig. 7 for the prompt, and Fig. 8 for an example response from ChatGPT o1-mini. The API to o1-mini does not include access to the tokens used for chain-of-thought, so this example is

[11] Tokens are counted with the tiktoken implementation of BPE tokenizer o200k_base: https://github.com/openai/tiktoken/tree/main.

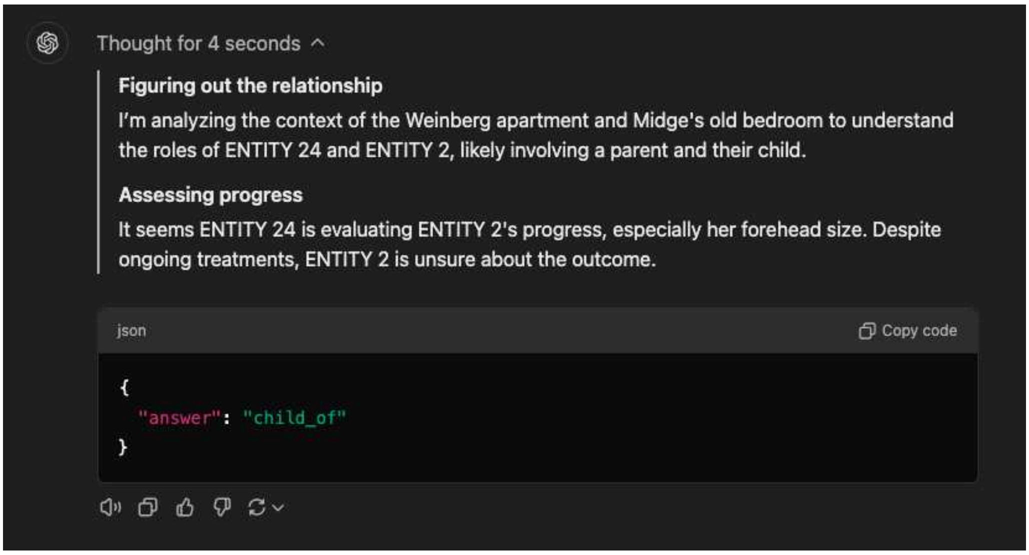
```
messages = [
    {
        "role": "system",
        "content": f"""Your goal is to extract relationships between TV characters in a
            scene of a TV series.
You will be provided with their dialogues, wrapped in <dialogue>.
Speaker names start with `ENTITY`, and their lines are separated by `:`.
You will read the dialogue and identify the relationship between a certain pair of
    entities, as requested in <question>.
The relationship is directed, so the order of entities in each triplet matters.
Here are the possible relationship types: {LABEL_OPTIONS}.
Here is an example:"""
    },
    {
        "role": "user",
        "content": f"""<dialogue>SCENE: INT. WEINBERG APARTMENT - MIDGE'S OLD BEDROOM -
            MOMENTS LATER

ENTITY 24: That forehead is not improving.

[ENTITY 24 lifts ESTHER out and lays her down on the bed.]

ENTITY 2: What? Are you sure?
ENTITY 24: It's getting bigger. The whole face will be out of proportion.
ENTITY 2: But look at her nose. It's elongating now, see?
ENTITY 24: The nose is not the problem. The nose you can fix. But this gigantic forehead
    ...
ENTITY 2: Well, there's always bangs.
ENTITY 24: I'm just afraid she's not a very pretty girl.
ENTITY 2: Mama, she's a baby.
ENTITY 24: I just want her to be happy. It's easier to be happy when you're pretty.
ENTITY 24: You're right. Bangs will help.</dialogue>
<question> ENTITY 2 is what of ENTITY 24? ANSWER with ONLY {LABEL_OPTIONS} </question>"""
    },
    {
        "role": "assistant",
        "content": "child_of"
    },
    {
        "role": "system",
        "content": f"""Great job! You have successfully identified the relationship
            between the two entities. Now, let's move on to the next one."""
    },
    {
        "role": "user",
        "content": f"""<dialogue>{scene_string}</dialogue>
<question> {head} is what of {tail}? ANSWER with ONLY: {LABEL_OPTIONS}.</question>"""
    },
]
```

Figure 6: Ones-shot prompt for LLaMA 3–70b–instruct.

included as a sanity check and see the thinking process of o1 models where the thought process is visible to us.

```
messages = [
    {
        "role": "system",
        "content": f"""You are a helpful assistant designed to extract relationships
            between TV characters in a scene of a TV series.
You will be provided with their dialogues, wrapped in <dialogue>.
Speaker names start with `ENTITY`, and their lines are separated by `:`.
You will read the dialogue and identify the relationship between a certain pair of
    entities, as requested in <question>.
The relationship is directed, so the order of entities in each triplet matters.

**Return only a JSON object** with the following property:

- "answer": one of the following {LABEL_OPTIONS}.

This property must always be present.

Do not include any additional text or explanations outside the JSON object.

<dialogue>SCENE: INT. WEINBERG APARTMENT - MIDGE'S OLD BEDROOM - MOMENTS LATER

ENTITY 24: That forehead is not improving.

[ENTITY 24 lifts ESTHER out and lays her down on the bed.]

ENTITY 2: What? Are you sure?
ENTITY 24: It's getting bigger. The whole face will be out of proportion.
ENTITY 2: But look at her nose. It's elongating now, see?
ENTITY 24: The nose is not the problem. The nose you can fix. But this gigantic forehead
    ...
ENTITY 2: Well, there's always bangs.
ENTITY 24: I'm just afraid she's not a very pretty girl.
ENTITY 2: Mama, she's a baby.
ENTITY 24: I just want her to be happy. It's easier to be happy when you're pretty.
ENTITY 24: You're right. Bangs will help.</dialogue>
<question> ENTITY 2 is what of ENTITY 24?</question>"""
    }
]
```

Figure 7: Prompt for OpenAI o1-mini.



Figuring out the relationship

I'm analyzing the context of the Weinberg apartment and Midge's old bedroom to understand the roles of ENTITY 24 and ENTITY 2, likely involving a parent and their child.

Assessing progress

It seems ENTITY 24 is evaluating ENTITY 2's progress, especially her forehead size. Despite ongoing treatments, ENTITY 2 is unsure about the outcome.

```json
{
  "answer": "child_of"
}
```

Figure 8: An example of prompting ChatGPT o1-mini.