



HAL
open science

Extraction and Analysis of Fictional Character Networks: A Survey

Vincent Labatut, Xavier Bost

► **To cite this version:**

Vincent Labatut, Xavier Bost. Extraction and Analysis of Fictional Character Networks: A Survey. ACM Computing Surveys, 2019, 52 (5), pp.89. 10.1145/3344548 . hal-02173918v1

HAL Id: hal-02173918

<https://hal.science/hal-02173918v1>

Submitted on 4 Jul 2019 (v1), last revised 31 Mar 2022 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Extraction and Analysis of Fictional Character Networks: A Survey

VINCENT LABATUT, Laboratoire Informatique d'Avignon – LIA EA 4128, France

XAVIER BOST, Orkis, France and Laboratoire Informatique d'Avignon – LIA EA 4128, France

A *character network* is a graph extracted from a narrative, in which vertices represent characters and edges correspond to interactions between them. A number of narrative-related problems can be addressed automatically through the analysis of character networks, such as summarization, classification, or role detection. Character networks are particularly relevant when considering *works of fictions* (e.g. novels, plays, movies, TV series), as their exploitation allows developing information retrieval and recommendation systems. However, works of fiction possess specific properties making these tasks harder.

This survey aims at presenting and organizing the scientific literature related to the extraction of character networks from works of fiction, as well as their analysis. We first describe the extraction process in a generic way, and explain how its constituting steps are implemented in practice, depending on the medium of the narrative, the goal of the network analysis, and other factors. We then review the descriptive tools used to characterize character networks, with a focus on the way they are interpreted in this context. We illustrate the relevance of character networks by also providing a review of applications derived from their analysis. Finally, we identify the limitations of the existing approaches, and the most promising perspectives.

Keywords: Information retrieval, Character network, Work of fiction, Narrative, Graph extraction, Graph analysis, Natural language processing, Multimedia processing, Image processing.

Cite as:

Vincent Labatut and Xavier Bost. 2019. Extraction and Analysis of Fictional Character Networks: A Survey . *ACM Computing Surveys* 0(0):0. <https://doi.org/0000001.0000001>

Note : This is a longer and slightly updated version of the official ACM CS article.

1 INTRODUCTION

The first works of fiction possibly date back to as far as the Paleolithic, and have constituted a major part of human culture since then [205]. Nowadays, it is estimated that, in average, adults are in contact with fictional stories for 6% of their time awake [205]. Besides their artistic and entertainment aspects, fictions are assumed to fulfill various social and psychological purposes, e.g. improvement of communication [132], development of empathy and collaboration skills [204, 224, 225], elaboration of social norms [204], proxy to understand the real world [243], assessment of social strategies [205], constitution of a collective memory [205]. It is therefore natural that they are abundantly studied by academia, and that fiction-related business is a significant part of the economy [13, 19, 164, 212].

A work of fiction takes the form of a *narrative*, i.e. a report of events telling a story. This report can be conducted through a variety of communication means: text, speech, image, music, gesture, and others, under a variety of forms: fables, tales, novels, plays, but also movies, TV series, video games, cartoons, and comics. The collection of events explicitly reported by the narrative constitutes its *plot*. These events are often ordered to form a chronological and/or causal chain [30]. By comparison,

Authors' addresses: Vincent Labatut, Laboratoire Informatique d'Avignon – LIA EA 4128, 339 chemin des Meinajaries, Agroparc BP 91228, Avignon cedex 9, 84911, France, vincent.labatut@univ-avignon.fr; Xavier Bost, Orkis, 610 rue Georges Claude, Pôle d'activité d'Aix-en-Provence, Aix-En-Provence, 13290, France, Laboratoire Informatique d'Avignon – LIA EA 4128, 339 chemin des Meinajaries, Agroparc BP 91228, Avignon cedex 9, 84911, France, xbost@orkis.com.

the *story* contains all the plot events, plus those imagined or inferred by the audience, based on both the plot and a number of contextual factors [30]. As an illustration, an *ellipsis* consists in removing events from the plot without affecting the story, as the audience will interpolate the missing parts. Put differently, the plot is *what* is told, whereas the narrative is *how* it is told, and the story is what the audience *perceives* of the plot through the narrative.

Historically, narratives have been studied from the Aristotelian perspective, which argues that the most important part of a narrative is its *plot*. However, more modern approaches focus on *characters* instead [15], and consider that they are the agents that advance the plot through their actions [167]. This is exemplified by Woloch in the field of *literary analysis* [271]. He defines the notion of *character-space* as the narrative environment of characters in a novel, i.e. their position relative to the other elements of the plot (place, time, other characters). In other words, this is how characters are described in the narrative. The concept of *character-system* extends this notion to the narrative as a whole, and corresponds to the union of all character-spaces. This approach has been noticeably used to study and understand how writers and directors build a narrative.

In addition to the characters themselves, researchers have then started to take into account the way characters *interact*, which is considered as the backbone of the narrative [55, 197]. In such a context, graphs are a natural modeling paradigm, as they allow representing and studying a system through the interactions of its constituting elements. A *character network* is a graph describing a narrative by representing the characters through its vertices, and the interactions between them through its edges. As we will see later, there are many methods to extract this type of network from some raw data representing the considered work, depending not only on the nature of these data, but also on the information that one wants to encode in the produced network, and on what one wants to do with it eventually. Moretti has shown that such an approach allows to handle more formally Woloch's concepts [170]. In a graph, the subgraph induced by a vertex and its neighborhood can be seen as a projection of the social aspects of the notion of character-space, whereas the whole graph, which contains all characters and their relations, represents the character-system [208]. Woloch emphasizes the fact that character-spaces must be considered jointly, and this is precisely what graphs, a naturally relational modeling framework, allow.

1.1 Value of Character Networks

This relevance of graphs for modeling works of fiction is illustrated by the number of articles dealing with character networks in the literature, and the variety of purposes for which they are used. We distinguish three categories of such articles.

First, in the context of *Narrative Analysis*, character networks are generally extracted manually, for a very small number of narratives (typically, a single one). Authors use them in a “distant reading” fashion to obtain a simplification of the plot [170], characterize the plot structure at various levels [209], detect relevant patterns and narrative events, identify character roles (e.g. protagonist vs. antagonist) or particularly important characters [208], assess the validity of literary theories [71, 75], and produce graphical representations [185, 259, 272]. In addition to the description of individual plots, they are also used to compare them, for example among episodes of a given series [97], or works belonging to the same genre [209], period [115] or author [200]. In other social science domains, character networks are also used for educational purposes [26], and to study certain psychological mechanisms [32].

Second, another category of works also adopts a descriptive and comparative approach, but relying on a *Complex Systems* paradigm. These authors consider that character networks are a type of *Complex Network*, and as such they apply the standard tools developed to analyze them [197], and/or propose new ones [53]. The network itself is the object of the study. Like for Narrative Analysis, these works generally consider a few narratives, as the networks are often extracted

manually. Many articles of this type compare the topological properties of character networks with those of other kinds of complex networks, e.g. real-world social networks [8, 154], random models [53], or other fictions [242].

Third, a large number of works originate from the *Artificial Intelligence* domain. They focus more on automating the network extraction process, which requires solving various text, speech, image, and/or video processing problems, depending on the media used. Compared to both other categories, this allows using much larger corpora. These works also consider character networks as models of the plot, and take advantage of this to solve higher-level problems: role detection [118], genre classification [11, 238], storyline detection [267], story segmentation [269], movie scene segmentation [148], video abstraction [252], recommendation systems [144], and others. The results obtained by solving some of these problems can be used to treat higher-level tasks, e.g. the detected roles can help summarize a plot. Certain authors directly relate character networks to novel fields such as *movie information retrieval* [191], which consists in obtaining and exploiting valuable information from collections of movies.

Character networks even reach the mainstream audience, mainly for their relevance as a visualization tool. Numerous non-academic or educational Web pages display character graphs extracted from popular culture works, such as *Star Wars* [22, 84, 85, 211], *Harry Potter* [134, 202], *Marvel* movies [123], *Love Actually* [203], *Game of Thrones* [59, 89, 106, 114], *Star Trek* [193], *The Simpsons* [198], *The Office* [183, 227], *Seinfeld* [233], *Curb your Enthusiasm* [76], *Grey's Anatomy* [150, 266], and *Friends* [9, 219, 223]; as well as from classics like *Sherlock Holmes* [39], European drama [80] and Shakespeare's plays [40, 96, 111].

1.2 Specific Features of Fiction Works

The extraction and use of character networks concern all types of works, including non-fictional ones. For instance, certain authors focus on biographies [258], professional meetings [86], journal articles [237], and broadcast news [260]. So in theory, it is possible to apply these methods developed for non-fiction to deal with fiction. However, in practice this does not necessary leads to good results, because works of fiction possess some specific features, absent from non-fiction. These can result in specific issues, whose resolution requires suitable processing, but they can also correspond to additional information one can leverage through appropriate methods to improve performance. We give examples of both aspects in the following.

First, there can be differences in the *structure* of the narrative. For instance, plays and TV or movie scripts are *semi-structured*, in the sense that scenes are explicitly bounded and speakers are explicitly named. This feature can be harnessed during network extraction [175], and does not appear in non-fictional narratives (or most other types of fictional ones, for that matters). In video-based narratives, the set of camera and editing rules, conventions and guidelines, sometimes metaphorically called *film grammar* [34], is quite different in fiction and non-fiction works. For instance, the so-called *180 degree* rule states that during a scene, the relative positions of the characters on the screen must not change. The *shot alternation* (or *shot/countershot*) rule is particularly used during conversations: it specifies that consecutive shots alternatively show the involved characters. Yeh *et al.* [273] leverage both of them to improve character detection in movies. Comics and animated films are apart, as in their cases, the medium itself is unlike anything related to real-life. Characters can be highly deformed human, non-anthropomorphic beings, or even inanimate objects, which makes ineffective the methods designed to detect faces or persons in photographs [13] or live action movies [256]. Moreover, the structure of comics narratives is unique, in the sense that they include information under a variety of forms encoded in both text (captions, speech balloons, onomatopoeia) and drawings (pose, graphical conventions) not found in other media, and whose extraction requires specific methods [13].

Second, there are generally significant *stylistic* differences. In texts, literary prose is considered as more complex than journalistic prose [70], and even more so when the work is older [98]. One of the effects of style is actually to give a unique identity to the work, and to distinguish it from both non-fictional works and other fictions [31]. Stylistic differences are so marked that it is possible to assign works automatically to their creators [11]. They significantly affect the performance of generic methods on a variety of NLP tasks: plot modeling, character detection and story generation [69], text summarization [120], named entity recognition (NER) [12, 255], co-reference resolution [130, 255]. There are a number of reasons for this drop in performance. For instance, for character detection in novels [69]: many characters are relatives and share the same last name; they bear nicknames; some fictional characters are inanimate objects in real life; writers use specific honorifics corresponding to complex, possibly outdated and even imaginary social conventions; and they craft names in order to convey certain meaning or function. In fact, this task is difficult even for humans, enough to require a specific annotation process [254]. For co-reference resolution, the problem comes from longer sentences, more frequent use of pronouns and direct speech, more numerous and shorter co-reference chains [130]. Similarly to text, certain characteristics of fiction works make generic audiovisual processing tools inefficient [273]. For instance, movie directors use a variety of complex, possibly genre-related, editing techniques [147]. At a lower level, the same face can appear under a variety of lights, colors, angles, expression, and other deformations, which do not correspond at all to the very controlled conditions under which non-fictional works are recorded (e.g. news or talk show). Speech-wise, conversations are subject to background noise or music, involve more participants, and a way of speaking that is unlike that found in other forms of audiovisual productions [41].

Third, fictions often are *closed-worlds*, in the sense that they are self-contained and involve recurring entities, possibly with made up names. Generic tools ignore this characteristic, which sometimes can help handling certain tasks [116] such as alias resolution (finding the different variants of a character's name). On the contrary, most generic tools rely either on a training corpus or on external databases: in both cases, the described entities are likely to be completely different from those occurring in a fiction. For example, a standard approach when performing face-matching in news is to leverage pictures from press articles and their captions: this cannot be done for movies containing fictional characters [282]. If anything, this method is more likely to return the actor's rather than the character's name. Similarly, many NER systems rely on gazetteers or services such as DBpedia, Wikidata, or YAGO [31], which are likely to include only the main characters (if any) of the considered fiction. For instance, none of the proper nouns used in Tolkien's *The Lord of the Rings* would be present in a standard list of first names or places.

Fourth, there is also a difference in the way characters interact in works of fiction, compared to real life [208]. A real-world social network represents an auto-organized system, whose structure emerges from the interactions between some agents acting according to their own agenda. By comparison, the writer or director controls all actions of fictional characters, and arrange them according to a plot. Put differently, real-world networks are the result of microscopic processes, whereas fictional ones are caused by a macroscopic process [208]. There is no reason to suppose that the writer tries to mimic actual social relationships when producing the work of fiction. As we will see later, studies show that this is generally not the case, as numerous character networks extracted from fictions do not exhibit realistic topological properties. This is because other constraints come into play, such as the intelligibility and appeal of the plot. Analyzing a different structure is likely to require specific tools, compared to real-world networks (including non-fictional character networks).

1.3 Perimeter and Organization of the Survey

The first publications related to fictional character networks date back to the early 2000s, e.g. [8, 232]. As explained before, both extracting and leveraging these networks involve solving specific problems. However, there is no synthetic review describing the solutions proposed in the literature. With this survey, we want to fill this gap. Not only do we consider articles directly related to fictional character network extraction and/or usage, but also articles focusing on certain specific steps of this process (without necessarily trying to deal with such networks). Note that certain authors extract other types of graphs (non-character-based) from works of fiction, such as scene transition graphs [276], or narrative structure graphs [117], but we do not include them in this review.

Our contributions include the description of the tools currently available and the approaches currently adopted to detect characters and their interactions from all forms of narratives, as well as the methods leveraging them to build character networks. We also contribute by identifying open problems at all levels of the extraction and analysis processes, and proposing perspectives to solve them.

Terminology-wise, we need to distinguish scientific work from work of fiction. For this purpose, we will use the words *author* and *article* to refer to scientific authors and their work, whereas *writer* (or *director*, *playwright*, or any medium-specific term) and simply *work* will refer to artistic authors and their works of fiction.

The rest of the survey is organized in two parts. We first focus on the process of extracting a character network from a work of fiction. We introduce it in a generic way (Section 2), before describing its three main steps: the identification of characters and their occurrences (Section 3), the detection of their interactions over the narrative (Section 4), and the extraction of the graph itself (Section 5). In the second part, we focus on how to leverage character networks. We first discuss the descriptive tools used in the literature to characterize them, and then examine a selection of more elaborate tools developed to solve specific problems (Section 6). Finally, we identify the current main issues of the field, and conclude with some perspectives (Section 7).

2 OVERVIEW OF THE EXTRACTION PROCESS

The process of extracting character networks from works of fiction depends a lot on the form of the considered narrative, e.g. novels are not treated like movies. In order to give the reader a general overview, in this section we make abstraction of these differences and present this process in a very generic way. In the rest of the survey, on the contrary, we focus on their differences.

We consider that this process consists of three main steps, represented in Figure 1: 1) the identification of characters; 2) of their interactions; and 3) the extraction of the proper graph. Each of them can be conducted in a number of ways, depending not only on the nature of the considered narrative, but also on the planned usage of the character network, and on certain methodological choices.

The first step is the most dependent on the form of the narrative, as it starts with the raw material, i.e. the work of fiction itself. We distinguish two substeps. The first is to detect occurrences of characters in the narrative, for instance looking for people names in a novel, or looking for faces in a movie. The second is to unify these occurrences, i.e. to determine which ones correspond to the same character. In a text, the same character can appear under different names, whereas in a movie, the same face can be shown under a variety of scales, colors, lights, and angles. The output of this step takes the form of a chronological sequence of unified character occurrences.

The second step consists in detecting interactions between characters. Note that it is sometimes more efficient or convenient to conduct parts of this process during the first step, but this is generally

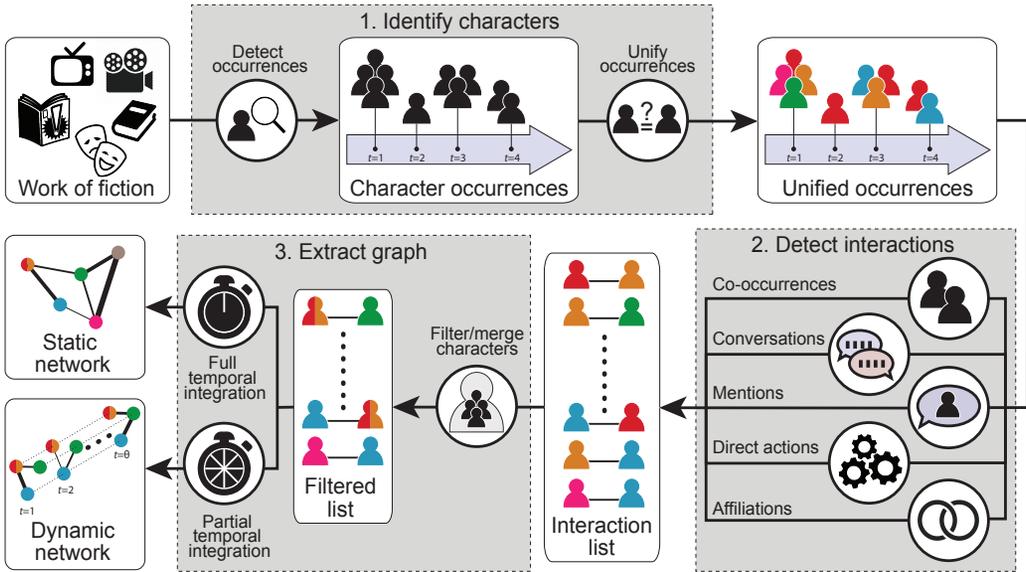


Fig. 1. Overview of the generic character network extraction process. Figure available at [10.6084/m9.figshare.7993040](https://doi.org/10.6084/m9.figshare.7993040) under CC-BY license.

not the case. We identify five different definitions for the notion of interaction. Many authors consider that a simple *co-occurrence* between two characters is enough to infer an interaction between them. Others prefer to identify *explicit* interactions, which is generally a more difficult process. One way of doing this is to take into account *conversations*, and to consider that two characters interact when one talks to the other. With certain forms of narrative such as plays, in which speakers are given, this task is relatively straightforward. An alternative is to focus on the content of the conversations, and to leverage *mentions*, i.e. situations where one character talks about the other. Some authors consider all sorts of actions one character can perform on the other (besides conversing). This is particularly the case with novels, a form of narrative in which such actions are explicitly described. Finally, certain authors do not focus on actions and prefer to use *affiliations*, i.e. explicit or inferred social relationships such as being married, being relatives, or working together. Note that it is possible to combine these definitions of the notion of interaction, for instance by looking for both co-occurrences and conversations.

The output of the second step is a chronological sequence of interactions between characters. The third step is therefore relatively generic, as it relies only on this list and is thus independent from the nature of the original narrative. We distinguish two substeps. The first, which is optional, consists in simplifying this sequence by *filtering and/or merging* some of the characters under certain conditions. For example, when considering co-occurrences, some authors merge characters that always appear together: this allows simplifying the network. The second substep defines how the graph is extracted through *temporal integration*, i.e. the aggregation of the previously identified interactions. There are a number of approaches for this purpose, which we separate into two groups: those performing a *full* integration and therefore leading to a *static* network, and those performing only a *partial* integration, and producing a *dynamic* network.

3 CHARACTER IDENTIFICATION

Character identification consists in detecting *which* characters appear in the considered narrative, and *when* exactly they appear in this narrative. As mentioned before, the form under which characters appear in the narrative varies much depending on the medium. In the case of text, they can be represented in three ways [70, 251]: *proper nouns* (e.g. “Sherlock Holmes”), *pronouns* (e.g. “He”), and *nominals*, i.e. anaphoric noun phrases referring to characters (e.g. “The consulting detective”). For videos, they either can appear onscreen, or be mentioned in the audio stream (again as a proper noun, pronoun, or nominal). In comics, characters can either appear as drawings or be mentioned in the text (again, under the same three forms).

Automating character identification is quite challenging, which explains why many non-specialists prefer to perform this task manually. We describe this manual approach separately, as there are various ways of proceeding (Section 3.1). But our focus is rather on automatic approaches, for which we distinguish two subproblems: first, finding character occurrences in the narrative (Section 3.2); and second, determining which of these occurrences represent the same character (Section 3.3). For both subproblems, there are two very different categories of approaches, which depend on whether characters are represented in a textual vs. audiovisual way. Note that this dichotomy does not necessarily matches the type of narrative, for instance a movie can be treated as a video or as a text (through its transcript or script).

Besides strict character identification, certain authors perform some additional processing in order to extract individual attributes to describe characters (e.g. age, gender...), and/or to filter them. We discuss this in Section 3.4.

3.1 Manual Approaches

Some authors adopt a fully manual approach to detect character occurrences, in which case there is no need to distinguish occurrence detection from occurrence unification, as both tasks are conducted at once.

3.1.1 Direct Annotation. The most widespread method is direct annotation, which consists for the authors in annotating by themselves the narrative they want to study, e.g. [32, 136] for novels, [154, 168] for myths, [170, 228] for plays, [55, 248] for movies, and [21, 267] for TV series. Opting for such a manual approach can be due to technical limitations, e.g. the authors do not have access to efficient automatic methods [170]. However, it can also be a methodological choice, e.g. to better focus on the assessment of the other steps of the network extraction and/or analysis process [3, 4, 21].

Though rarely mentioned explicitly [97], the authors that adopt manual approaches for later extracting *co-occurrence* networks (cf. Section 4.1) often ignore mentions of characters which are just named by others, but do not physically participate in the action, as in [197]. It is also not clear exactly how annotators deal with occurrence unification. However, context generally suggests they perform this task, and do so manually (e.g. [170]), as the extra cost is marginal (cognitively speaking).

3.1.2 Character Index. Instead of doing the annotation work themselves, certain authors take advantage of predefined resources, which are also manually constituted. For certain classic novels, literary experts have constituted so-called character indexes, indicating at which point of the plot each character appears. This is for instance the case for Rousseau’s *Les confessions* in [206–208], and Park’s *Toji* in [188, 189]. Several authors proceed similarly for comics [8, 92], as they study the

Marvel universe by taking advantage of the *Marvel Chronology Project*¹, an online database listing the occurrences of all significant Marvel characters.

Even if such indices are elaborated by experts of the considered work of fiction, it is difficult, or even impossible to assess their reliability. Moreover, it is important to notice that they impose a predefined level of precision on the rest of the extraction process. For instance, character occurrences are expressed in terms of pages for *Les confessions*, and comic issues for the *Marvel universe*. This lack of control can be considered as a limitation, since the level of precision affects certain subsequent extraction steps (e.g. it constrains the selection of a narrative unit when extracting co-occurrence networks, cf. Section 4.1).

Like for direct annotation, the elaboration of indices is likely to include some form of character occurrence unification. However, it is difficult to determine whether it is the case for a given index, and according to which procedure exactly. Indeed, this task is conducted by the writers of the index, not those of the study that take advantage of this index for network extraction. Moreover, the index is often not properly documented regarding this aspect. Only a very few articles mention occurrence unification, but they nevertheless reveal some differences in the way they handle this task. For instance, in [206, 208], the character index considers all variants of the character names, but not pronominal references, whereas the index used in [207] includes both.

3.1.3 Crowdsourcing. In practice, it is hard to handle more than a few works of fiction when using either of the previous approaches (direct annotation vs. predefined character index). A workaround is to turn to *crowdsourcing*, as Rochat & Kaplan do in [209] to constitute their own indices. Interestingly, this study is also characterized by its multimedia nature, as the authors consider a corpus of science-fiction works including novels, comics, movies, TV series, and video games. The manual approach has the advantage of allowing a more uniform network extraction process over the variety of considered media, and therefore makes it possible to compare them. They can also select their own level of precision during the elaboration of the indices: pages for novels and comics, and intervals of one minute for movies and TV series. For video games, they experiment with three different base materials: walk-throughs (i.e. texts explaining how to finish the game), which are treated like novels; cinematic scenes, which are treated like other videos (movies and TV series), and transcriptions of these scenes, which are treated like scripts.

3.2 Detection of Character Occurrences

We now switch to approaches that are at least partially automated. As mentioned before, the process of character identification largely differs depending on whether the narrative is visual (Section 3.2.3) or textual, which is why we separate them in our description. Moreover, certain texts such as plays and scripts possess a structure which can be leveraged for character occurrence detection, so we distinguish such semi-structured text (Section 3.2.2) from free text (Section 3.2.1).

3.2.1 Free Text. As mentioned before, a character can appear under three forms in text: proper noun, nominal, and pronoun. The methods used in the literature all handle the first form, but not necessarily the two others, as detecting them is generally a much harder problem, and they are often not considered as informative. A simple way to detect character names is to use a predefined list of these names and proceed through exact matching [23, 29, 99, 110, 139]. Such a list is generally constituted manually, either by the authors themselves or through an external source such as the Wikipedia page of the considered novel. Constituting it is not a trivial task, as characters can be referred to through a variety of *aliases*, i.e. variations of their name. For instance, Sherlock Holmes

¹<http://www.chronologyproject.com/>

can also be called simply “Sherlock”, “Holmes”, or “Mr. Holmes”. Some authors perform a manual verification after the exact matching step [99].

Detecting character names can be viewed as a specific version of the Named Entity Recognition (NER) problem. NER consists in finding expressions in the text corresponding to proper nouns [174], and to identify their category (e.g. *Location*, *Person*, *Organization*). A number of authors apply off-the-shelf NER tools to novels, e.g. [5, 48, 71, 105, 229], and then only retain the *Person* entities. Incidentally, those are generally much more frequent than other proper nouns in literary texts [257]. Dekker *et al.* perform an empirical comparison of four such tools in the context of character extraction based on novels [61]. It is possible for a NER tool to assign different categories to distinct instances of the *same* string, because of contextual differences. For instance, “France” is a country, but also a first name. However, in the context of novels, such a situation is likely an error: it is generally agreed upon that a novel is a small, self-contained world [11], and that the writer would not confuse the reader by using the same name to denote entities of different types (such as a person and a place) [12]. A straightforward solution to the multiple category issue is then to keep the majority category, as in [11, 12]. Valls-Vargas *et al.* specifically train a classifier to distinguish characters from other types of mentions [256, 256]. A very few authors consider other categories in addition to *Person*, such as *Location* [140] and *Organization* [10], which eventually results in a network with multiple types of vertices.

Fiction texts have certain characteristics which are leveraged by some authors, either to perform some post-processing after having applied an off-the-shelf NER tool, in order to find missed mentions and/or discard incorrectly detected ones, or to design new fiction-specific NER tools. A simple method is to remove infrequent names, as they are likely to be errors. For instance, Elsner [69] and Sack [214] remove names appearing fewer than five times. Some authors also perform a manual verification to fix the errors of the automatic tools [214]. Some automatic approaches use honorifics (titles such as “Sir” or “Madam”), generally by relying on manually predefined resources. One can take advantage of a list of honorifics to detect them in the text and check the surrounding text for character names [11, 12], or look for a set of patterns describing the various possible combinations of honorifics, first names and last names [251]. Some approaches proceed similarly with action verbs, as only characters are likely to be their subjects. For instance, Ardanuy & Sporleder [11] use a manually constituted list of speech verbs (e.g. to say, to discuss), while Goh *et al.* leverage WordNet to focus on *human* action verbs only [93]. Zhang *et al.* also use the grammatical structure of the sentence through part-of-speech (PoS) tagging [281]. Finally, some approaches consist in looking for relations of possession (through genitive marks, such as “s” in English) [251], as only characters are supposed to own things. These approaches are more robust than generic NER tools, in the sense that they allow detecting non-human characters behaving as humans [119]. In [31], Bornet & Kaplan propose an ensemble-based method associating most of the aspects listed above, for French texts. It relies on the vote-based combination of the outputs of six basic classifiers. Each one focuses on the detection of a specific type of clues: presence of honorifics, position in the sentence, semantics of neighboring words, grammatical structure, occurrence in external resources, and presence of nearby explicit quotes.

These last methods are likely to return not only proper nouns, but also nominals (anaphoric noun phrases referring to characters). Some authors propose methods specifically designed to detect these nominals, generally through regular expression matching. Elson *et al.* [72] look for structures of the form: a *determiner* (article, possessive, number...), an optional *modifier* (e.g. an adjective), and a *head noun* (not necessarily a proper noun). They manually compile lists of determiners and head nouns based on their corpus and external linguistic resources such as WordNet. They use them to detect the determiner and head noun first, and consider the text located in between as the modifier.

The task of detecting pronouns is more or less difficult depending on the considered language. For English, exact matching based on a manually defined list is a simple and efficient approach [70].

3.2.2 Semi-Structured Text. A number of narratives can take the form of a script: theateric plays, movies, TV series. A script is essentially a conversation-based text, with specific structure and formatting described as *semi-regular* by certain authors [2]: scene boundaries are clearly indicated, the characters involved in a scene are explicitly listed at its beginning in uppercase, and the name of the character speaking a line is indicated right before it, also in uppercase. When the script is properly formatted and structured, it is relatively straightforward to extract this information automatically. Authors have proposed methods based on exact string matching [63, 118, 173], regular expression [238, 272], or a custom parser [200].

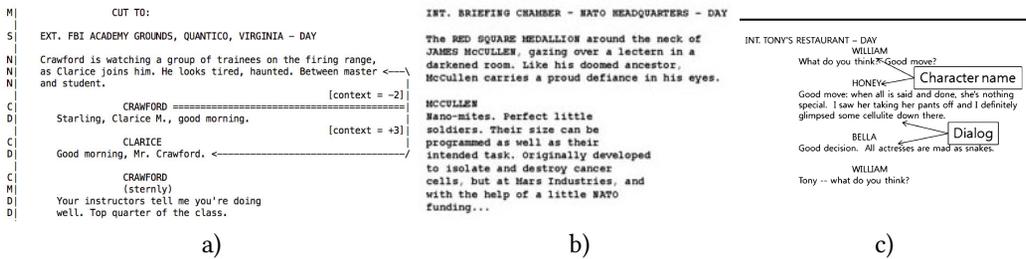


Fig. 2. Examples of semi-structured text: a) from [2], b) from [63] c) from [118].

However, this structure and formatting is not a proper standard, and can vary from one script to the other [160], as illustrated in Figure 2. It is even possible to find inconsistencies in the same script. Machine learning can help solving this issue. Agarwal *et al.* [2] propose a method to identify which parts of the script correspond to character lists, dialogues, speaker names, scene boundaries, and scene instructions. Tan *et al.* [242] take advantage of this type of decomposition to only focus on character mentions associated to utterances, in order to ignore passive characters which are present in a scene but do not intervene.

Discrepancies can also appear in the speaker names. In this case, a simple approach consists in using an *a priori* list of the characters involved in the script [129, 242, 272], with their associated aliases. This list is generally constituted manually, or by taking advantage of publicly available resources (generally also constituted manually), such as the Wikipedia page of the considered work of fiction. Again, machine learning based method can be more robust than such simple matching-based approaches. Makris & Vikatos [160] take advantage of the Wikipedia pages of the movies they study to train a classifier into identifying which character speaks which line. Certain authors directly apply off-the-shelf NER tools [94, 282] to detect speaker names.

Identifying the characters involved or speaking during a scene is enough when extracting co-occurrence (cf. Section 4.1) or conversational networks (Section 4.2), respectively. Dealing with other types of interactions between characters requires identifying character mentions in the rest of the text [135, 173]: not only explicitly identified speakers, but also scene metadata, spoken lines, and/or stage directions. In this case, one can apply similar approaches to those already described for free text. For instance, Krishnan & Eisenstein [129] train a classifier to detect addressees mentioned in utterances, in order to determine who speaks to whom exactly. They detect not only proper nouns, but also nominals, including titles and placeholder names (e.g. “bro”, “dude”, “sir”).

Choi *et al.* [53] apply an approach relatively similar to those designed for scripts, but rather to a biographic dictionary of fictional characters: the classic *Dictionary of Greek and Roman mythology*

by Grant & Hazel. This work takes the form of a series of entries, each one summarizing the biography of a given character. Choi *et al.* constitute the character list by parsing the entry keys, and identify their occurrences in the entry bodies through exact matching.

3.2.3 Visual Narratives. In audiovisual narratives, detecting character occurrences amounts to solving several distinct but related problems, depending on whether one focuses on the video or the audio stream. In videos, these are face detection and face tracking [273, 274].

Video Streams. *Face detection* consists in identifying which parts of a still image correspond to faces, as illustrated in Figure 3. See [133] for a recent review of the field. Jung *et al.* note that current face detection methods are efficient mainly on front views of the faces [118]: this is a strong limitation in our context, as a character can be filmed under a variety of angles. Weng *et al.* also observe that current automatic methods do not reach satisfying enough performances, which is why they first proceed manually [267]. However, they later train their own model to obtain acceptable performances on their dataset [268, 269]. They experimentally find the community structure of the extracted networks to be relatively robust to face detection errors. A number of authors proceed automatically using off-the-shelf tools [148, 177, 190, 252, 273].

The face detection problem is relatively similar when dealing with comics, except that the images are drawings (cf. Figure 3). This implies a number of additional difficulties: the characters can be very deformed, non-human, or even non-anthropomorphic. Moreover, the structural lines defining the characters and objects composing the panels are mixed with textures, screentones, and stylistic elements. For these reasons, methods designed to handle photographs generally perform poorly on comics, which require specific approaches [54, 235]. One such approach consists in adapting features or models originally developed for photographs, e.g. facial landmarks (points corresponding to specific parts of the face such as eyes or mouth) in [235]. Takayama *et al.* [241] handcraft features to fit the specific case of mangas (skin and hair colors, jaw line shape, symmetry). The fact that a pattern appears frequently in the narrative is also used as a hint to distinguish characters from other objects [107]. More recent articles focus on training Deep Neural Networks [54], but there is not enough publicly annotated data yet to reach the full potential of such approaches [13]. Finally, it is worth noticing that before being applied, many face detection approaches require some preprocessing, in particular detecting panel bounds and speech bubbles [54], which in turn constitute specific problems [201, 234].

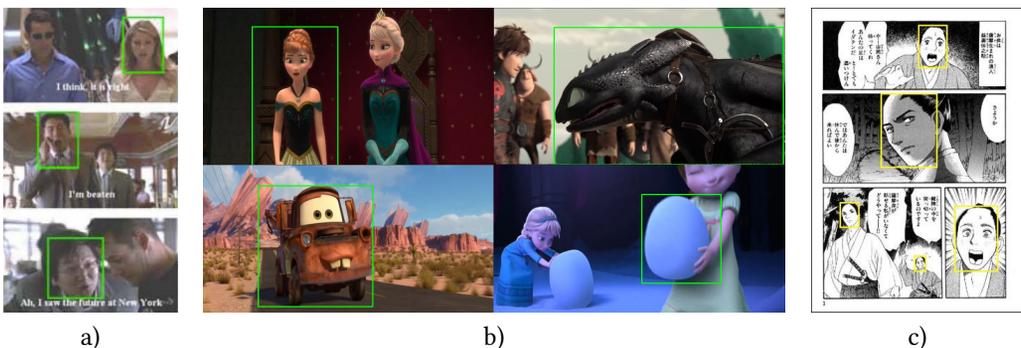


Fig. 3. Examples of face and body detection in: a) live-action movies [190]; b) animated movies [226]; c) comics [54].

Face tracking builds upon face detection, and aims at identifying chronological sequences of faces corresponding to the same person in a video. These sequences are called *face tracks*, and can

be considered as character occurrences in videos. Performing face tracking requires accounting for changes in pose, scale, rotation, expression, color, light, angle, and blur. Most authors use off-the-shelf tools [148, 282], usually based on some form of similarity-based classification of the detected faces.

Somandepalli *et al.* [226] detect characters in *animated* movies. This task proves to be much more difficult than with live-action videos, as the design of the characters can vary widely, including non-human, and even non-anthropomorphic shapes (cf. Figure 3). The authors first list character candidates by detecting salient objects in a generic way, before taking advantage of graphical and saliency features to discard irrelevant ones. They then use an off-the-shelf tool to track deformable objects.

Audio Streams. When using the audio stream, detecting character occurrences amounts to solving the *speaker segmentation* (or *speaker change detection*) problem [169]. It consists in partitioning the audio stream into segments associated to unique speakers. Put differently, one wants to find the moments corresponding to switches between speakers. This task is sometimes performed simultaneously with that of *segment clustering*, which consists in grouping the segments spoken by the same person. Performing both these tasks sequentially or simultaneously is called *speaker diarization*. However, we treat this later in Section 3.3.2, and focus here only on speaker segmentation.

Certain existing systems work well in controlled environments, but this performance strongly drops when applied to fiction works, e.g. movie trailers and cartoons [57], and TV series [73]. This is mainly due to the presence of background music and sound effects, the higher number of speakers [35], the spontaneous (though acted) nature of the exchanges, and the shorter speech turns [41]. Results improve when using methods specifically designed or trained on fictional audiovisual narratives, e.g. [18] for movies, [37] for TV series. It is worth noting that compared to video-based methods, audio-only tools do not allow identifying characters that appear in a scene *without speaking* [118].

Multimodal Approaches. Certain approaches try to combine several types of information, be it video-, audio-, or language-based. A few multimodal methods able to perform speaker segmentation using both audio and video have been proposed, but we describe them later as they all additionally solve speaker clustering (and therefore speaker diarization). Scripts can be used to distinguish speakers or on-screen characters as in [118, 190, 191]. A script is not time-stamped, so this approach requires first solving an additional problem called *script alignment*, which consists in determining the exact time at which each line contained in the script occurs in the video. In [51], Chen *et al.* use transcripts extracted from TV series. They apply off-the-shelf tools to detect all three forms of textual character mentions (proper nouns, nominals, pronouns).

3.3 Unification of Character Occurrences

The second step of character identification is occurrence unification, which consists in determining, for each detected occurrence, to which character it corresponds. Like for occurrence detection, the methods proposed for this purpose vary much depending on the medium of the narrative: textual (Section 3.3.1) vs. visual (Section 3.3.2). However, this time there is no distinction to make between free and semi-structured text, as all the additional information of the latter has already been used during the detection step.

As the literature shows, character unification is often not performed at all. There are mainly two reasons for this: this task is generally harder than occurrence detection (especially in text); and in certain situations it is simply unnecessary. For instance, when extracting a purely conversational network (Section 4.2) from a clean script, the speakers explicitly named in the script are enough, e.g. [63, 170, 190]. Or when extracting a network from a novel by considering chapter co-occurrences

(Section 4.1) [126]: it is likely that all characters participating in the chapter will be explicitly named. Some authors show this empirically, e.g. Seo *et al.* [222] argue that restricting their analysis to explicit proper nouns (and thus, ignoring pronouns and noun phrases) is enough to perform their targeted tasks (character ranking and edge prediction) without significant performance loss.

3.3.1 Textual Narratives. As mentioned before, characters occurrences appear under three forms in text: proper nouns, nominals, and pronouns. Unifying these occurrences can be considered as a specific version of the *coreference resolution* problem, which consists in identifying sequences of expressions, called *coreference chains*, that represent the same concept (see [239] for a recent review). Generic tools exist to solve this problem, but their performance does not necessarily translate to fiction works [255]. In particular, they tend to overlook minor characters such as those mentioned only through nominals (e.g. “the detective”) [255]. Moreover, in our case the referents are necessarily persons (characters), a category of entities possessing certain characteristics (e.g. gender) which can be leveraged to improve performance. Two variants of the problem appear in the literature: certain authors focus only on *alias resolution* (e.g. [71]), which consists in grouping proper nouns referring to the same character, while others additionally solve pronominal and/or nominal anaphoras.

The task of alias resolution arises because of the variability of proper nouns appearing in fiction works. On the one hand, in addition to their full name, characters are generally called by a variety of aliases depending on context, style, and other factors. For instance, Sherlock Holmes can also be called “Mr. Holmes” or “Sherlock”. On the other hand, some aliases cannot be unequivocally associated to a character, e.g. “Mr. Holmes” can refer to both Sherlock Holmes and his brother Mycroft Holmes. Most authors use some form of *name clustering* to perform alias resolution, each cluster corresponding to all the names encountered for a specific character. Roughly, They use two factors to determine that two aliases point at the same character: *string similarity* and *gender compatibility* [11, 69, 72, 185, 255]. The gender of a character mention can be detected using gendered honorifics (e.g. “Mr.” vs. “Mrs.”) and gendered first names (e.g. “Stephen” vs. “Stephanie”), matched to a manually constituted list or some external resource such as WordNet [72, 185].

A straightforward approach to compare strings is to use an appropriate distance function [113]. However, by doing so, one ignores the structure of the names: potential presence of honorifics, initials, multiple first or last names, distinction between first and last names. Moreover, a number of conventions are culture-specific (e.g. the use of patronyms in Russian names). Also, the relative proportions of first and last name occurrences is likely to vary considerably from one work to the other, as it is tied to stylistic aspects: it is assumed to reflect the level of intimacy in the narrative [257]. Certain authors propose to perform direct comparisons through predefined patterns [185] or rules [11, 12]. Elsner [69] first compares only multiword names, in order to deal with the ambiguity of isolated first or last names. He constitutes clusters of similar and compatible multiword names, and then only assigns single word names to these clusters whenever possible. The remaining names are assigned based on spatial proximity in the text and lexical frequency. Certain authors use a generative approach [71, 255]: based on multiword names found in the text, they produce potential variants thanks to predefined recombination rules (e.g. addition of honorifics, omission of first names) and resources such as gazetteers. These artificial names are then matched to those found in the text. Vala *et al.* [255] use additional constraints to prevent certain names from being grouped together: co-occurring names, names with the same last name but different first names, names containing different honorifics.

The other types of anaphoras are more difficult to handle, as they convey additional issues. Some pronouns or nominals may not be connected to any proper noun (and therefore character), if their referent is missing. They can also have split referents, e.g. “They” and “the Holmes brothers”

can both refer to “Sherlock and Mycroft Holmes”. Certain anaphoric expressions can also refer to non-character entities. Many authors use off-the-shelf tools to solve automatically pronominal anaphoras, e.g. [229, 236, 251]. Lee & Yeung additionally define a distance limit between the reference and the referent, in order to discard relations deemed too remote [140]. Vala *et al.* [255] extend to pronouns and nominals the cluster-based approach used for alias resolution. As for proper nouns, gender compatibility can be leveraged for certain pronouns (ex. “she” vs. “he”) and nominals (ex. “uncle” vs. “aunt”). In order to identify anaphoras referring to characters (by opposition to other types of entities), they constitute a list of verb-noun co-occurrences considered as frequent in novels, and perform a grammatical dependency parsing: only the expressions involved in such situations are considered as character mentions. In [113], Jannidis *et al.* use a co-reference resolution tool that they previously developed specifically for novels in German [130]. In particular, they use linguistic resources to associate close synonym nominals to the same character.

3.3.2 Visual Narratives. Like before, handling audiovisual narratives amounts to solving very different problems, which depend on whether one uses video or audio data. When dealing with videos, the problems at hand are face track clustering, and possibly face-name matching.

Face track clustering consists in identifying groups of face tracks (output from the occurrence detection step) corresponding to the same face, hence character. Certain authors use off-the-shelf tools [148, 252, 273], but others consider that these generic methods are not sufficiently efficient when applied to fiction works [267]. Their results can be improved by training them on a corpus of such works [268, 269], but this requires additional work and resources. Zhang *et al.* [282] propose a new method combining the Earth-mover’s distance with constrained k -means clustering, followed by an additional pruning step.

As mentioned before, when using the audio stream, the problem is to perform *speaker clustering* based on the output of the speaker segmentation step, the whole process being called *speaker diarization* [169]. Speaker clustering consists in grouping audio segments spoken by the same character (akin to track face clustering). Generic methods are subject to the same limitations as observed for speaker segmentation [41, 57, 73]. Methods specifically developed for fictions obtain better results. In [37], Bost & Linares treat TV series through a two-step method first solving the problem locally at the scene level, then combining these partial results at the global level in order to deal with the whole character set. In [38], they turn to a multimodal approach to enhance their method: in addition to their audio-based tool, they independently perform speaker diarization based on low-level video and audio features, before performing optimal matching to combine both resulting outputs. However, after further experimentation, Bost *et al.* consider the obtained performance is not sufficient, and eventually turn to manual annotation [35, 36].

Like for character occurrence detection, certain authors adopt multimodal approaches. Some prefer to extract the transcript of the work and apply text-based methods instead of directly using the video or audio stream. This is the case for Chen *et al.* [51] who, after having detected character occurrences in this text, use existing off-the-shelf co-reference tools to detect chains of mentions to the same character. They identify the concerned character by using predefined rules, exploiting the presence of a proper noun in the chain, or connections to utterances whose speaker could be identified. They propose an automatic method based on agglomerative convolutional networks to take advantage of the latter type of information [52] when solving co-references and identifying characters associated to co-reference chains. In [41], Bredin & Gelly combine face track clustering and speaker diarization: they first detect speakers through standard speech activity detection tools, before using face embeddings to cluster the face tracks corresponding to the resulting speech segments.

An additional issue specific to audiovisual narratives is to determine the name of the characters detected by grouping faces or speech segments. In the former case, this is called *Face-name matching*, and in the latter, *speaker identification*. In both cases, solving this problem requires using linguistic information: speech content (via transcripts, scripts, or subtitles), text overlaid in the video, predefined list of characters, other external resources. For instance, based on the assumption that the title of the fiction work is known, Tran *et al.* [245] retrieve its list of characters from IMDb, look for their picture using Google Image, and leverage this information to infer character names in the movie through matching. In the context of fiction works though, the favored approach is to leverage scripts [118, 148, 190, 282]. This again raises the issue of *script alignment* (with the corresponding video, transcript, or subtitle), as scripts are not time-stamped. After alignment, the script directly allows recovering speaker names, and inferring addressee names (for instance by crosschecking names mentioned in the conversation and on-screen faces).

For comics, a variety of methods have been proposed for *character detection* (or *face recognition* [241]), i.e. to match the multiple occurrences of a character's face. The general approach consists in defining some form of similarity measure, which is then leveraged to group occurrences corresponding to the same character. This is what Takayama *et al.* [241] do, based on the features they use for face detection (cf. Section 3.2.3). Stricker *et al.* [235] adopt a similar approach, but to compare their sets of facial landmarks. The problem is difficult and still open, its resolution will likely require large annotated corpora [13]. Ho *et al.* [107] represent character occurrences by graphs of adjacent graphical subregions, and use approximate graph matching to group occurrences corresponding to the same objects. Sun *et al.* [240] propose a method based on local features.

3.4 Additional Processing

In addition to character occurrence detection and unification, certain authors perform some additional processing related to character identification. We identify two such operations (which are not mutually exclusive): filtering characters (or character occurrences) considered to amount to noise (Section 3.4.1), and get a more detailed individual description of the characters (Section 3.4.2).

3.4.1 Character Filtering. Certain authors remove characters deemed not frequent enough [69, 71, 110, 236]. For instance, Elson *et al.* [71] remove a character if it appears three times or fewer in a novel, or if he amounts to 1% or less of all occurrences. They consider these as noise, possibly generated by their alias resolution method: the occurrences are likely to refer to another existing character, and not a separate infrequent one. When extracting conversational networks, Suen *et al.* [238] ignore a character if he speaks fewer than five times over the plot. Character filtering can concern a large part of the novel, e.g. Hutchinson *et al.* [110] remove approximately 80% of the characters to study the *Harry Potter* novels.

Bossaert & Meidert [32] remove characters depending on their social category. Indeed, they want to use the character network of the *Harry Potter* series of novels as a proxy to study peer support among adolescents. For this reason, they remove all characters that are not Harry Potter's schoolmates.

3.4.2 Attribute Extraction. In addition to simply identifying the characters, certain authors extract some additional information allowing to describe them. It typically takes the form of nodal attributes in the final character network. These are retrieved directly from the work of fiction itself, but also from external sources.

The most popular character trait used in the literature is probably their gender, which can be assessed manually [32, 213], through some external resource [92], or automatically [11, 12, 72]. Some authors define a categorical attribute representing the various sides present in the considered narrative; in the articles reviewed in this survey, this is always performed manually. For instance,

when dealing with the *Marvel Universe*, Gleiser extracts the characters' *alignment* (*hero vs. villain*). In [277], Yose *et al.* study the medieval Irish epic text *Cogadh Gaedhel re Gallaibh*, which narrates a war between Irishmen and Vikings. For their analysis, they distinguish three categories of vertices: *Irishmen*, *Vikings*, and *Others*. When studying Homer's *Illiad*, Kydros *et al.* identify four mutually exclusive categories of characters: *Greek*, *Trojans*, *Gods* and *Others*.

Certain authors focus on more sociological traits. In [209], Rochat & Triclot are interested in the representation and relative position of science and politics in science-fiction works. Through crowdsourcing, they manually define an attribute representing whether the main occupation of each character is mainly related to *Politics*, *Technology*, or *Science*, but also *Family*, *Religion*, or *Art*. They use the same attribute to identify *Animals*, and the remaining characters are noted as *Undefined*. When studying the *Harry Potter* series of novels, Bossaert & Meidert manually retrieve the school house and school year of each student (in addition to his or her gender) [32]. In [213], Rydberg studies Greek tragedies. He manually extracts the so-called social class of attributes: a mix between race, social status and narrative role (*Gods*, *Upper/lower class mortals*, *Chorus*, and *Non-speaking characters*).

The resulting nodal attributes have mainly two usages. Certain authors use them directly to study the considered narrative. For instance, Bossaert & Meidert want to assess the effect of certain character traits on the psychological mechanism of peer support, and therefore on the character network structure [32]. Other authors take advantage of these attributes to solve some higher-level problem. For instance, Ardanuy & Sporleder [11, 12] use character gender to build features later leveraged to classify novels automatically.

4 INTERACTION DETECTION

Based on the character occurrences, the next step of the extraction process consists in detecting all interactions happening in the narrative between each pair of characters. Such an interaction can be explicitly described, but also inferred from the narrative, depending on what one considers to be an interaction. We identify five distinct approaches in the literature.

The first (Section 4.1) is co-occurrence-based, and relies on a decomposition of the narrative into smaller narrative units. Two characters are considered to interact when they *jointly appear* in the same such unit. The second approach (Section 4.2) considers only *direct verbal interactions* between the characters. This is particularly appropriate for dialogue-oriented narratives such as plays. The third (Section 4.3) requires one character to *explicitly mention* another one to infer an interaction between them. The fourth (Section 4.4) takes into account other types of *direct interaction* than conversation (e.g. fighting, kissing). The fifth (Section 4.5) focuses on explicitly expressed *affiliations*, such as family relationships or being coworkers. Finally, it is also possible to combine several of these approaches, in various ways (Section 4.6).

4.1 Co-occurrences

The co-occurrence-based approach is the most widespread in the literature, probably because it is the easiest to apply: detecting interactions in a more precise way can be a difficult problem, even for humans [209]. This approach consists in breaking down the considered work into smaller *narrative units*, and in assuming that two characters interact when they occur together within the same unit. A few authors use additional constraints, to ensure that co-occurrences actually capture interactions. Some want the narrative unit to contain *only* the two characters of interest, and no one else [69, 112]. Others take into account only *consecutive* occurrences, i.e. not separated by another character [98].

Using co-occurrences presents several limitations mainly caused by their imprecise nature. Indeed, co-occurrence is only a proxy for actual interaction, as it is possible for two characters to

appear together without interacting at all (e.g. they both are spectators of some event [167], or one mentions the other in his absence [197]). The first limitation is that this imprecision propagates to the network itself: the set of co-occurrence-based interactions theoretically contains the conversation-, mention-, action- and affiliation-based ones, plus some false positives². In practice though, Ardanuy & Sporleder [12] argue that false positives are rare in the sense that two co-occurring characters are almost always related in one way or the other. But this holds only for their experimental results, obtained by integrating co-occurrences over whole narratives. On the contrary, Edwards *et al.* note that co-occurrence networks are denser [68].

The second limitation also directly comes from the imprecise nature of co-occurrences. As they encompass a number of different types of interactions, it is not possible to assign them a direction, and they are therefore regarded as some form of bilateral interaction. Furthermore, for Kwon & Shim [135], due to their imprecise nature, co-occurrences ignore intimate aspects of interactions, such as opinions and emotions. The third limitation, according to Prado *et al.* [197], is that using co-occurrences results in more importance being given to otherwise minor characters, when later analyzing the obtained character network. As discussed in our introduction, all these arguments must be balanced accordingly to the possibly very specific nature of the considered narrative.

We discuss the choice of the narrative unit in Section 4.1.1, as it depends on the type of narrative and can affect the end result. Besides the detection of interactions under the form of co-occurrences, certain authors additionally assign a numerical score to each interaction in order to include more information in the character network eventually extracted: we review such approaches in Section 4.1.2.

4.1.1 Narrative Unit.

Novels. In novels, Rochat & Kaplan use the page as a narrative unit [208], as imposed by the predefined character index they leverage during character identification (see Subsection 3.1). Such a partitioning of the text, based on purely *physical* (and therefore arbitrary) aspects, results in the possible split of chapters, paragraphs, or even sentences. It is therefore very likely to miss co-occurrences. Rochat & Kaplan try to overcome this problem through a two-page unit with a one-page overlap (instead of consecutive pairs of pages).

Other authors use smaller narrative units that can be considered as more natural, in the sense that they at least avoid such arbitrary splitting: one sentence [140], ten sentences [189], one paragraph [69], ten paragraphs [71]. However, the length of sentences and paragraphs can vary considerably from one writer to the other, to the point where one writer's paragraph can be shorter than an other's sentence [112]. Using word spans instead solves this problem: the literature contains examples ranging from five [110] to 1,000 words [99]. Certain authors take chapter breaks into account, so the span can be smaller when reaching the end of a chapter [99]. However, word spans are as arbitrary a narrative unit as the page, and suffer from the same limitation.

Certain authors prefer to use a larger unit, especially the chapter [12, 126, 167], or the pericope [97] (its counterpart in the context of biblical writings). Like the sentence and the paragraph, however, it does not lead to split sentences but its size can vary significantly from one writer to the other. Moreover, it can be considered as too long a narrative unit for this usage, as many events can take place in the same chapter. Bolioli *et al.* [26] use what they call a *narrative sequence*: a small portion of text characterized by its unity of location and involved characters (basically, a scene). Prado *et al.* [197] proceed similarly when studying *La chanson de Roland*, as this song of heroic deeds is originally divided into stanzas of very irregular size.

²False negatives are also a possibility, but they require the characters to interact without co-occurrence or mention, which is highly unlikely.

Scripts. For scripts, certain authors use the line [272], but the most widespread narrative unit is simply the scene, e.g. [63, 94, 200, 238]. However, as noted by Suen *et al.* [238], if a script contains long scenes, this can lead to the connection of characters involved in completely different parts of the story. To solve this issue, Stiller *et al.* [231, 232] prefer to use *subscenes*, i.e. parts of a scene identified by the fixed set of involved characters. Put differently, as soon as one character leaves or enters the stage, the current subscene ends and a new one begins.

Comics. In comics, Rochat uses a two-page narrative unit [209], like he does for novels (see above). In [8, 92], the authors use the comic book issue, but this narrative unit is dictated by the predefined database they use as raw material. There seems to be a lot of room for exploration here, as comic book formats are very diverse, and there are very few articles dealing with this question. For instance, a seemingly natural choice would be to transpose the concept of scene, and consider a sequence of panels involving the same characters as a narrative unit. However, this requires efficiently solving certain lower-level problems, in particular panel identification and panel ordering. Existing methods to detect the boundaries of panels take advantage of the black lines generally outlining them, or of the white space called *gutter* separating them [13, 234]. But a number of artists use complex page layouts, which makes both panel detection and ordering much harder: overlapping panels, panels joined by other objects (speech bubbles, caption) partly open panels, or even panels with no explicit boundary [234].

Videos. In videos, certain authors use a fixed time interval as a narrative unit, e.g. ten seconds [144], one minute [209]. Similarly to using the page as a narrative unit for novels, this approach seems quite arbitrary. Indeed, it is likely to split the plot at any moment, including for instance the middle of an action. Several more *natural* units are traditionally used when dealing with videos [191, 210]: The *frame* is the smallest, it is a single image. The *shot* is a sequence of frames continuously recorded by a single camera. The *scene* is a collection of consecutive and semantically related shots. The *group* is a logical subdivision of a scene (e.g. a conversation, in a scene containing several of them [190, 191]), and the *sequence* is a series of consecutive scenes constituting a short story.

A number of authors prefer to work with scenes, but this requires solving the difficult problem of scene boundary detection. Certain authors perform this task manually, e.g. Weng *et al.* in [267]. They later adopt a semi-automatic approach, consisting in applying first an existing automatic tool and then manually detecting and correcting its errors [268, 269]. Other proceed automatically, using an off-the-shelf tool [252, 282]. In [244], Tran *et al.* segment movies by detecting stable periods in terms of character on-screen presence, and merging them according to criteria related to character similarity and temporal proximity.

4.1.2 Interaction Score.

Positive Scores. While most authors adopt a boolean approach when detecting interactions (i.e. presence vs. absence), certain authors try to assess their *intensity* under the form of numerical scores. Some use the distance between the two concerned character occurrences, expressed as a number of some smaller narrative unit separating them (or, more exactly, some decreasing function of this quantity). For instance, Park *et al.* [189] use a narrative unit of ten sentences, and experiment with two distance functions: the numbers of words and sentences separating the two occurrences. Grayson *et al.* [98] count the number of tokens between occurrences.

Others use the duration of the co-occurrence, e.g. the number of lines spoken in the scene [238], or the number of seconds during which both characters jointly appear on screen in a video [244, 248]. Rieck & Leitte [200] use the proportion of words spoken by the considered pair of characters, relative to the total number of words spoken during the scene. This amounts to giving more importance to

the most talkative characters. Yeh *et al.* [273] take advantage of film-editing guidelines, especially the *180 degrees* and *shot alternation* rules (cf. Section 1.2): their score is the number of consecutive shots showing alternatively both concerned characters.

Signed Scores. In addition to the intensity of an interaction, a few authors try to detect its *polarity*, i.e. whether or not it is friendly or hostile, by leveraging the context of the co-occurrence. This results in a *signed* score (i.e. it can be negative). Some authors adopt an *indirect top-down* approach consisting in assessing the general polarity of the concerned narrative unit, and extending it to all interactions occurring in it. For novels, Min & Park apply an off-the-shelf tool to perform sentiment analysis and measure the polarity of each narrative unit constituting the plot [165]. For movies, Ding & Yilmaz train a support vector regressor into estimating the polarity of a scene based on a set of visual and auditory features [63]. The same authors also experiment with affinity learning to take advantage of *visual concepts*, a higher-level information describing the scene *context* through the objects and environments it involves [64].

Alternatively, the *indirect bottom-up* approach consists in assessing the emotional state of the individual characters, instead of the whole narrative unit, and obtaining the polarity of an interaction by combining the states of the concerned characters at the time it occurs. Lee *et al.* do so for movies [144]: they use off-the-shelf tools (completed by manual correction) to assess the emotional states of characters based on both conversational (movie script) and visual (facial expressions) cues. In comics, a similar approach could be applied by leveraging graphical elements such as effect lines, characters' facial expressions and poses, and conventional symbols used to reflect emotions. However, this is not possible yet, as only a handful of articles propose methods to detect such elements [13].

Finally, certain authors propose a *direct* way to monitor the polarity of an interaction. For novels, Chaturvedi *et al.* [49] use a Markov model to represent the chronology of the interactions between a given pair of characters, and detect friendly vs. hostile phases. For this purpose, they leverage a collection of features taking into account the content of the text associated to the co-occurrences (e.g. type of action, structure of the sentence, grammatical role of the characters, lexical aspects). This model is itself embedded in a semi-supervised framework allowing to train over a partially annotated corpus of novel summaries. They later propose an extended version of their tool in which the nature of the interactions is not necessarily binary, and is learned in an unsupervised way [47, 48]. In [112], Iyyer *et al.* use a relatively similar principle, except their model is a neural network, which is able to learn relationship descriptors in an unsupervised way. Instead of a single *friendly/hostile* score, each co-occurrence is therefore represented by a set of descriptor values, each one with its own semantics (e.g. *love, violence, education*).

4.2 Conversations

A number of authors focus only on verbal interactions between two characters, i.e. when one explicitly *talks* to another, in order to later extract so-called *conversational networks*. Technically, the process of detecting them is generally harder than for co-occurrences, as one has to detect that a conversation is taking place, as well as to distinguish the involved speakers and addressees. A fundamental difference between co-occurrence and verbal interaction is that the latter is naturally *unilateral*: one character talks, the other listens. This generally leads to the extraction of *directed* networks [191].

As explained before, verbal interactions can be assumed to be subsumed by co-occurrences, as two characters need to co-occur in order to converse, but can co-occur without necessarily talking to each other [35]. One could therefore suppose that only focusing on conversations leads to some information loss. However, certain authors argue that this is not the case, as many aspects

of interpersonal relationships [102], if not most [264], are conveyed through conversation. The validity of this argument actually depends on the type of the considered narrative, or even on the specific narrative itself. Several authors note that conversational networks are suitable only for narratives rich in verbal interactions [140, 165]. Lee & Yeung [140] give the counterexample of the *Book of Genesis*, in which the tense relationship between Abraham’s wife and servant is mentioned frequently in the text, while they never speak to each other. Ardanuy & Sporleder [12] identify plays as the most appropriate form of narrative for conversational networks, as they are essentially scripted dialogue, by comparison with novels, in which most of the action takes place off-dialogue (e.g. McCarthy’s *The Road*, Yourcenar’s *Mémoires d’Hadrien*). Visual media such as comics tend to rely only lightly on dialogues.

As detecting verbal interactions generally requires identifying *who* talks to *whom*, the appropriate method depends on the nature of the narrative. Like before, we distinguish textual (Sections 4.2.1 and visual narratives (Section 4.2.2). We then turn to methods proposed to estimate the intensity and direction of verbal interactions (Section 4.2.3).

4.2.1 Textual Narratives. In text, one can distinguish utterances (quotations of a character) from proper narration parts (description of the action occurring). Verbal interactions can take one of two forms: *direct* speech, which consists in explicitly quoting the utterance, and *indirect* speech, which consists in reporting what the speaker said (e.g. “Arthur told his knights to go home”). Conversational networks are built upon the former, whereas the latter belongs to the class of direct interaction considered in Section 4.4.

Identifying verbal interactions in fictional free text is a difficult task. First, the text generally contains much more proper narration parts than utterances, and both are tightly intertwined (e.g. speech verbs connecting utterances during a conversation). Second, the typographical conventions used to identify utterances in the text are not always respected (not all spoken text is quoted), and are sometimes ambiguous (e.g. sneer quotes). Third, the speakers are often not identified explicitly [90], and when they are, this information can be embedded in the narration. Therefore, automating the detection of verbal interactions requires the resolution of several NLP-related problems: detecting utterances, assigning a speaker to each of them, and possibly also identifying its listeners. Maybe for this reason, certain authors decide to proceed manually [75].

By comparison, semi-structured text is easier to handle, as typography generally allows distinguishing utterances from the rest of the text, and speakers are explicitly indicated. Thus, the only remaining problem is to identify addressees. As for novels, this task is difficult to automate, which is why certain authors perform it manually [170, 172, 213].

Quote Detection. It is relatively straightforward to detect quotes when the text is clean, by simply relying on the presence of quotation marks [45, 71, 264] and/or using simple rules [275]. It is worth noting that some complications can appear, depending on the file encoding: for example, Unicode includes tens of distinct glyphs likely to be used as quotation marks.

But quotes are not always used correctly, or the considered language can be without typographic quoting convention, e.g. old Chinese punctuation style [139]. In this case, the authors rely on the presence of quotative verbs, and on the sentence structure. Zhang *et al.* [281] train a decision tree based on morphological and typographical features, in order to recognize quoted speech and detect speaker changes between two consecutive utterances. Mamede & Chaleira [161] propose twelve heuristics to identify both direct and indirect utterances in Portuguese. These are based on typographical clues, verb tense, presence of certain pronouns, temporal adverbs, and interjections.

Quote Attribution. The literature contains a number of articles dealing with quoted speech attribution for general texts [70], but we focus on the methods proposed specifically to handle

fictions. For a given utterance, one can distinguish two cases: either the speakers are mentioned explicitly (be it by their name or some anaphora), or they are implicit and can be inferred from the context.

The most widespread method to deal with *explicit* speakers relies on detecting the speech verb associated to the concerned utterance, and considering its subject as the speaker. Certain authors propose to manually define [91, 102, 264] or automatically learn appropriate rules [90], other look for patterns [130, 139]. In the absence of a speech verb, certain authors proceed similarly with nearby action verbs [161], whereas others look for the nearest character occurrence (excluding the utterance itself) [281].

The first step in identifying *implicit* speakers is generally to select a subset of candidates among all known characters. For this purpose, one first has to estimate the boundaries of the conversation: this can be done by leveraging long proper narration parts separating sequences of utterances. For this purpose, Waumans *et al.* [264] use empirically determined distance and length thresholds, and also include the narration located right before and after the conversation. The candidates correspond to the characters present in the obtained portion of text.

A number of authors then deal with implicit speakers by assuming that the characters involved in the considered conversation respect certain rules relative to conversation turns (a.k.a. the *conversation turn assumption*) such as: consecutive utterances are not spoken by the same character, or one speaker answers to the previous one. It is then possible to leverage the explicit speakers identified before to infer the missing ones. Certain authors identify speaker alternation patterns and use them to define rules [130, 139, 264], or to learn them in a supervised way [90]. He *et al.* translate these patterns into features describing the position and frequency of character occurrences relative to the considered utterance. These features (as well as Celikyilmaz *et al.*'s ACTM, see below) are then used to train a classifier [102]. Some authors additionally use gender compatibility to improve utterance attribution [102].

Elson & McKeown train a classifier [71, 72] to deal with implicit and explicit speakers at once. They use standard low-level features as well as the syntactic category of the utterance. The latter is related to its structure and context in the paragraph containing it (e.g. being followed by a speech verb and a character name). They also rely on the *conversation turn assumption*, as the classifier is aware of the last identified speaker. O'Keefe *et al.* [186] note that this last point is not a realistic assumption, and consider instead the problem as a sequence labeling one. Using the same features as Elson & McKeown, they experiment with various sequence decoding methods.

Certain authors focus on the utterance content, in order to distinguish speakers based on the topic they talk about. Chaganty & Muzny [46] propose a supervised method based on neural networks and use GloVe vectors to represent words [192], but they do not obtain conclusive results. Celikyilmaz *et al.* [45] build an Actor-Topic model (ACTM), a Bayesian model allowing to predict the probability that a given character spoke a given utterance, depending on the topics mentioned in this utterance and those mentioned before by this character (and the others). Each utterance is assigned its most probable character according to this model.

Addressee Identification. The addressee is sometimes mentioned *explicitly* in the utterance, when the speaker directly calls him. Certain authors define specific rules to take advantage of this situation [173]. Krishnan & Eisenstein [129], in particular, focus on the level of formality of these verbal interactions. He *et al.* [102] train a classifier to detect these cases, based on features related to punctuation and typical forms of interjection.

When the addressee is *implicit*, one uses the *conversation turn assumption* already leveraged for speaker identification. Certain authors define rules assuming that a speaker talks to the preceding speaker in the conversation [135, 175], or the few preceding speakers [238], and/or to the following

one [163, 173, 195]. In the case of conversations involving more than two characters, certain authors assume that a speaker talks to everyone present [36, 79, 173], which can also be considered as co-occurrence with an additional constraint (not just being present, but also speaking).

The two-stepped supervised method proposed by Yeung & Lee [275] applies a similar conversation turn-based principle to assign simultaneously both speakers and addressees to utterances. They first train a CRF-based classifier to identify explicit speakers and listeners, using morphological, grammatical and positional features. They then train another CRF classifier to identify the conversation boundaries, and use the *conversation turn assumption* to fill the missing speakers and listeners from the first step.

4.2.2 Visual Narratives.

Comics. We could not find any article dealing with the extraction of conversational networks from comics, mainly because this requires solving a number of lower-level problems for which there are not much literature [201] and no efficient solutions yet [13]. The first ones are panel detection and ordering, already mentioned in Section 4.1.1. Then, it is necessary to detect speech bubbles and captions. This problem is difficult, because the shape and position of bubbles widely vary depending on the artist, culture, context, meaning, and a number of other factors. Existing approaches range from the exploitation of low-level color-based features (e.g. detecting white blobs) to adaptive outline detection methods [201].

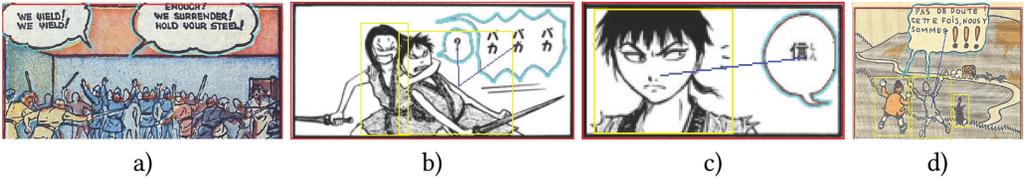


Fig. 4. Examples of distinct cases of speech bubble use in comics [201]: a) unidentifiable speakers; b) in-between speakers; c) out-of-panel speaker; d) multiple speakers.

One also needs to retrieve the shape of the bubbles, as it generally conveys a specific meaning: smooth for speech, cloudy for thoughts, spiky for screams. But there is no definitive convention, for instance certain artists use captions to show the characters' thoughts. The location of the bubble is also important. First, relative to the other bubbles, it directly affects the reading order, and is needed to extract properly a conversation. Second, relative to the characters, it indicates who speaks. Bubbles often possess a tail directed at their speaker, which helps solving the speaker attribution problem. However many different cases exist, see Figure 4 for instance. Rigaud *et al.* [201] adopt a graph-based approach leveraging the distance between bubbles and characters, and the angle of the tail. But again, using a tail is not a universal convention, which makes it harder to perform this task in general.

Finally, the last step is optical character recognition (OCR). Comics can be very challenging for standard OCR models [13], for a number of reasons: 1) many authors write by hand; 2) use of a variety of fonts, sizes, and colors, sometimes in the same bubble, sometimes mixed with pictographs; 3) the environment is very noisy due to the drawings surrounding the text, which is sometimes integrated to the background (e.g. sound effects). In conclusion, there are a number of open low-level problems to solve before being able to extract conversational networks from comics.

Videos. There are not many articles describing the extraction of conversational networks based on audiovisual narratives either. The problems are roughly the same as for the other media: detecting utterances, identifying their speaker, and possibly their addressees.

When using the video stream, one has to detect which parts of a face track correspond to a speaking character. Certain authors use off-the-shelf tools to identify lip motion [190, 282], which allows simultaneously detecting utterances and assigning them to speakers. When using the audio stream, the speaker diarization tools described in Section 3.3.2 also solve the two same problems simultaneously. In theory, it is also possible to apply Automatic Speech Recognition tools to the audio stream in order to obtain a transcription and get access to the utterance content, or to use directly the script if it is available, however in practice we did not find any use of these approaches in the fictional network extraction literature.

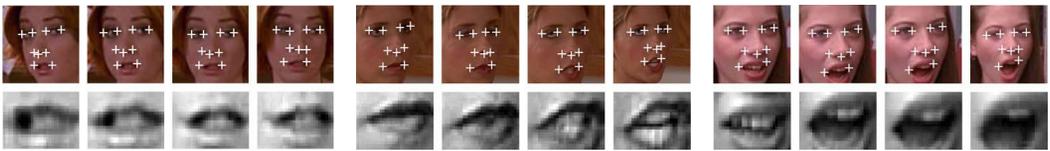


Fig. 5. Examples of lip motion detection for three different characters of the TV series *Buffy the Vampire Slayer* [74].

Regarding the identification of addressee, one approach is to use the *conversation turn assumption* that we already mentioned for text. For TV series, for instance, Bost *et al.* [35, 36] consider that a speaker talks to another character if the latter speaks before and after the considered utterance. For an utterance starting or ending the conversation, they just consider the following or preceding speaker. When there are more than two speakers involved in the conversation, they use temporal proximity to determine the addressees. Alternatively, it is also possible to take advantage of lip motion detection: one can consider that characters appearing without moving their lips are addressees [191]. However, this approach has the major drawback of missing addressees that are not shown on-screen.

4.2.3 Score and Direction.

Score. The ways scores are handled for verbal interactions are quite similar to those already described for co-occurrences (Section 4.1.2). Some authors adopt a Boolean approach, only identifying whether a verbal interaction is occurring or not. This can be for the sake of simplicity [173], or by lack of technical resources [170]. A number of authors use a numerical score to represent the intensity of each interaction. The simplest approach is to rely on the length of the interaction, be it expressed in words in the case of text [113, 238], or some temporal unit in the case of videos [35, 190]. Some authors prefer to compute the distance separating the utterances of the two concerned users, expressed in numbers of utterances, and apply a decreasing function, e.g. linear [238] or stepwise [195].

Like for co-occurrences, certain authors take advantage of the content of the conversation to compute signed scores reflecting the polarity of the exchange. In plays, Nalisnick & Baird [175, 176] use an existing lexicon designed for sentiment analysis, in order to associate a polarity (positive or negative) to certain words. The score of an utterance is the sum of these values over all of its constituting words.

Direction. Certain authors consider verbal interactions as symmetric [71, 170, 173, 264], in the sense that both the speaker and the addressee are interchangeable. It is an important simplification

though, as the information flows from the speaker towards the addressee, making the interaction asymmetric by nature: many authors prefer to consider that the speaker acts on the addressee [135, 139, 172, 190]. As we have seen before, the addressee is generally much more difficult to identify than the speaker, which could explain this choice. Certain authors go further in the simplification, and consider that all characters participating in a conversation speak to each other [264], which is conceptually very close to simply detecting co-occurrences.

Interestingly, Park *et al.* [190, 191] consider that a verbal interaction can involve a *single* character as both the speaker and the addressee, in case of *soliloquy* (i.e. when one talks to oneself). In terms of graphs, this results in loops, or self-edges (an edge connecting a vertex to itself). Note that such graphs are relatively unusual in the domain of complex network analysis, and most descriptive tools are not designed to take advantage of this information.

4.3 Mentions

Certain authors focus only on explicit mentions of characters during conversations. Compared to the conversational approach described in Section 4.2, the difference is that they consider that there is an interaction not when one character speaks *to* another, but when he speaks *about* another. Thus, like for conversational interactions, this requires detecting utterances and their speakers, but not their addressees. Also like them, this approach is appropriate for speech-oriented media. Once the utterances are identified, one only needs to list which character occurs within them: this is trivial, since character identification is the object of the first step of the extraction process (Section 3). This is the approach adopted to study novels, e.g. [71] (as a baseline, the main focus being conversational interactions), and movie [94] or TV [62] scripts. The approach from [53] also uniquely relies on mentions, but these are non-verbal: this is a very specific case, as the studied work is a biographical dictionary. The authors consider that there is an interaction between the character described by a given dictionary entry and all the characters mentioned in the body of this entry.

As with other types of interactions, it is possible to compute a score to characterize the intensity of a mention-based interaction, although most authors do not in this case. Elson *et al.* [71] count the number of times the character is mentioned in the concerned utterance, and divide by its length. Implicitly, they consider that the interaction is more intense if the utterance is short. Kwon & Shim [135] take advantage of the utterance content to compute signed scores in scripts. They look for utterances expressing the opinion of the speaker regarding the mentioned character. They use an off-the shelf sentiment analysis tool to detect the polarity of the utterance, and perform a manual pass to improve the result.

As a mention is by nature an asymmetric interaction, in the sense that some character performs an action involving another one [53, 62, 94], most authors consider them as such. However, a few authors choose to ignore this information [71, 173], either because they do not consider it as useful for the task at hand, or to simplify the processing.

4.4 Direct Actions

Instead of focusing only on verbal interactions, certain authors take into account all forms of direct actions that one character can perform on another (e.g. thinking about someone), or that two characters can perform jointly (e.g. fighting). By comparison with conversational interactions, detecting *general* actions allows handling narratives in which most of the interactions takes place off-dialogue (see Section 4.2). However, the task is even more difficult, as one needs to identify the action taking place, as well as the characters performing and undergoing it.

Some authors dealing with textual narrative proceed manually, possibly because they focus on certain classes of actions, semantically speaking: one character supporting another [32], physical

encounters [108], actions that influence the development of the story [259]. In [55], Cipresso & Riva are interested in identifying the interactions between movie characters according to four basic emotions: anger, fear, sadness and joy. For this purpose, they rely on a survey conducted over a sample of eleven persons watching the movie and detecting these interactions manually. When dealing with the *Game of Thrones* TV series, Stavanja & Klemen [230] focus only on the action of murdering someone, in order to predict who the next kill will be.

The automatic approaches proposed in the literature are all designed for textual narratives. Indeed, processing visual narratives implies recognizing actions in videos or images, which is a difficult problem called or *human-human interaction detection* [262]. Generic methods exist for videos [196, 262, 265], some of which have been applied to works of fiction (mainly movies). However, the literature does not seem to contain any article leveraging them to extract character networks. For comics, the problem can be considered as even harder due to the static nature of the medium: state of the art works only focus on the lower-level task of *pose detection* [13] (cf. Figure 6).

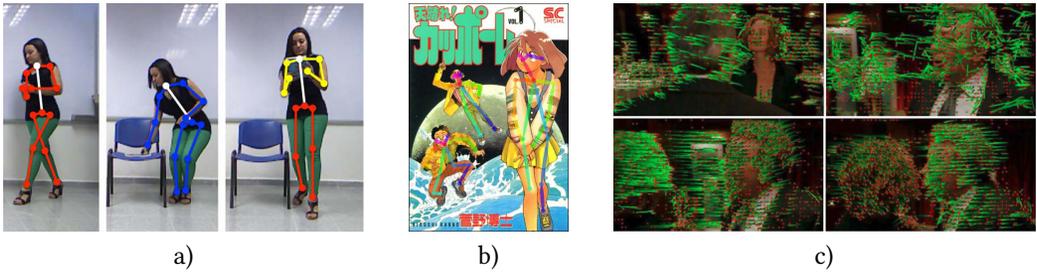


Fig. 6. Examples of identification of: a) poses in videos [149]; b) poses in comics [13]; c) human motion in videos [262].

Most automatic methods for text first employ parsing techniques to detect so-called SVO triplets (subject-verb-object), before filtering them in order to retain only those involving characters as subjects and objects [229, 251]. Certain authors take advantage of external resources such as *FrameNet* [14] to additionally filter the remaining triplets by focusing on certain classes of verbs, semantically speaking. To summarize, *FrameNet* is a database containing *frame semantics*, i.e. structured representations of situations involving various agents and objects, and semantic relationships between them (e.g. generalization) allowing to perform inference. Certain authors use predefined classes of interest, e.g. social interactions [140], while other estimate them *ad hoc*, e.g. the fifty most frequent [236].

In [1, 3, 5], Agarwal *et al.* detect interactions using their own tool, described in [6]. It is based on a classifier fed with essentially grammatical features, which allow it to detect how verb-related connections exist between certain characters. This tool was originally designed and trained to handle newswire text, so Agarwal *et al.* have to manually control its output when applied to novels. They treat a conversation as an interaction, so their interaction network subsumes the conversational network, as shown experimentally in [115].

Nijila & Kala [182] propose a supervised approach, not relying on SVO triplets but on simple co-occurrence of characters in the same sentence. They train a Convolutional Neural Network (CNN) to predict the class of action associated to each sentence involving at two characters. Of course, this method requires an annotated corpus, which they constitute manually.

Certain types of actions can be considered as naturally *unilateral*, as one character acts on another one, whereas other are *bilateral*, as several characters act together. In terms of graph extraction, this means extracting directed vs. undirected edges. Certain authors consider only unilateral [236]

or bilateral actions [140], and ignore the rest. Others use both, but consider all of them as bilateral, in order to avoid the identification of the subject and object characters, and therefore significantly ease the process . [251]. Agarwal *et al.* [3] distinguish between two types of actions: unilateral ones in which the object character is not aware of the action occurring (e.g. thinking about someone), which they call *observations*, vs. all other actions. They handle them separately, in order to compare them relative to the task at hand.

A very few authors compute a score to represent the intensity of the interaction. Trovati & Brady manually identify three categories of verbs [251]: *friendly*, *hostile*, or *unknown*. They ignore interactions described by verbs belonging to the later, and assign a signed score to the others, depending on the category. Srivastava *et al.* [229] use sentiment analysis to take advantage of the semantics of the textual context, and associate a polarity to the interaction (*cooperative* vs. *adversarial*).

4.5 Affiliations

Unlike the other types of interactions listed in this document, affiliations do not correspond to actions but rather to states. Among others, affiliations include: being blood-related, being married, and belonging to the same social group. They are explicitly mentioned in the narrative, which means that they can appear either in the conversation or in the narration. It is clearly the less popular approach, as only a few authors leverage this type of interactions. Moreover, the methods proposed in the literature only concern textual narratives.

Srivastava *et al.* [229] leverage an *a priori* selection of keywords (e.g. father, wife) to identify family relationships explicitly mentioned in the text. Krug *et al.* [131] propose a semi-supervised method to train a maximum entropy classifier in detecting family, romantic, professional and social relationships in literary texts. Starting from a set of annotated instances, it uses uncertainty-based active learning to select appropriate examples during training and ask the user to label them.

In [140], Lee & Yeung detect affiliations using a heuristic approach, based on a dependency parsing and the linguistic resource *FrameNet* [14]. They distinguish two cases: the relations stated *directly* (“Sherlock is the brother of Mycroft”), and the ones mentioned *in passing* (“Watson was referring to Sherlock’s brother Mycroft”). The authors treat both of them by looking for predefined patterns, designed using a development dataset. More specifically, they look for sentences involving several characters and fitting one of the manually selected semantic frames. Kokkinakis & Malm [127] apply a relatively similar template-based approach to Swedish novels.

Barbosa, Kondrak *et al.* [102, 159] adopt an utterance-based approach to extract family relations from novels. After having assigned each utterance to a speaker and a listener, they take advantage of family relationships explicitly mentioned in the text (e.g. MOTHER, DID YOU KNOW?). They then infer additional relations based on manually defined rules. For instance, if two female characters have the same parents, the system can infer that they are sister.

4.6 Hybrid Approaches

The five different definitions of the notion of interaction presented in this section are not mutually exclusive, and can be combined in various ways. For instance, Kyrios *et al.* consider all types of interactions except mentions and affiliations [136]; Mutton [173] and Deleris *et al.* [62] combine conversations and mentions; Lee & Yeung detect conversations, affiliations, and direct actions [140]. Srivastava *et al.* use co-occurrences, affiliations, and direct actions, and they also consider how similarly characters are described in the text [229]. When studying Edgeworth’s *The Absentee*, Falk *et al.* detect not only oral conversations, but also *written* ones [75], which belong to the class of direct actions in our nomenclature. He *et al.* use two distinct definitions, but in conjunction rather than disjunction: they use affiliations only if expressed in conversations [102].

Conducting manually the extraction process can be very costly, but this provides some advantages in return: some consider the obtained results as more reliable compared to automatic tools [21, 154], and this allows more flexibility when defining what an interaction is. Certain authors leverage this flexibility to simultaneously consider all forms of interactions [136, 168, 228] or distinguish between friendly and hostile interactions [154, 155, 157]. Others define additional constraints, e.g. focusing only on face-to-face interactions [197], or visual and physical contacts [21].

Mixing several types of interactions raises the problem of combining their laterality, as some may be unilateral and others bilateral. Kydros & Anastasiadis [136] proceed by replacing bilateral interactions such as co-occurrences by two distinct unilateral interactions involving the same characters, but possessing opposed directions.

The literature is lacking proper studies comparing the various interaction definitions, through their effect on the extracted network and possibly the resolution of the problem at hand. Jayannavar *et al.* [115] study various combinations: Agarwal *et al.*'s observations (cf. Section 4.4) only, other direct actions only, both forms of direct actions, and finally conversations as in [71]. They conclude the conversational interactions are indeed a subset of the more general ones. In [68], Edwards *et al.* compare networks based on manually identified direct actions with networks based on automatically detected co-occurrences, as well as conversations and mentions. They conclude that networks obtained automatically are good approximations with regards to density, centrality, and weight distribution, but not transitivity (see Section 6.1 for a description of these topological measures in the context of character networks).

5 GRAPH EXTRACTION

At this stage of the process, the characters and their interactions have been identified. The last step consists in building the character network based on this information. This requires making two methodological choices: how to define vertices and edges.

Vertices. In almost all the approaches presented in the literature, characters are represented as individual vertices. However, a few authors alternatively consider certain characters *collectively*, and represent them as a single vertex. This can be due to several reasons. The first is that some groups of people are mentioned in an indistinguishable way in the raw material, for instance: townfolks in Twain's *Huckleberry Finn* [126], Olympian gods and Greek soldiers in Homer's *Iliad* [259], by-passers in certain plays [137, 172] and novels [75]. Second, it can also be that some characters always appear simultaneously in the plot: considering them collectively can be viewed as some form of simplification. However, this can have an effect on many of the topological measures computed to describe the network, unless the multiplicity of the vertex is encoded in some way (e.g. vertex weight) and used when processing said measures. Third, it is possible that one wants to model the narrative at the level of groups rather than individuals, as Liu & Albergante [151] do for the *Game of Thrones* TV series, when they consider what they call *meta-entities* (houses, armies, religious groups) instead of characters.

Handling such collective vertices is difficult when their composition evolves over time. Moreover, it is possible for the same character to appear in the work of fiction both as an individual and as a part of a group. For these reasons, most authors choose to simply ignore such groups altogether, as they are generally secondary [3, 137].

Edges. As illustrated by Figure 7, the methods proposed to define edges differ much more from one author to the other, which is why we focus on them in this section. One has to consider several aspects of the interactions, which must be translated in graph-related concepts: their laterality, score, polarity, and temporality. The general approach is to represent unilateral interactions by directed edges, and bilateral ones either by undirected edges or by pairs of reciprocal directed

edges. The scores computed by certain authors to measure interaction intensity are modeled by edge weights, which can be signed to represent the polarity of the interaction (friendly vs. hostile). This results in graphs which can be (un)directed, (un)weighted and (un)signed. A signed network has both *positive* and *negative* edges, and is used to model antagonistic relationships [104].

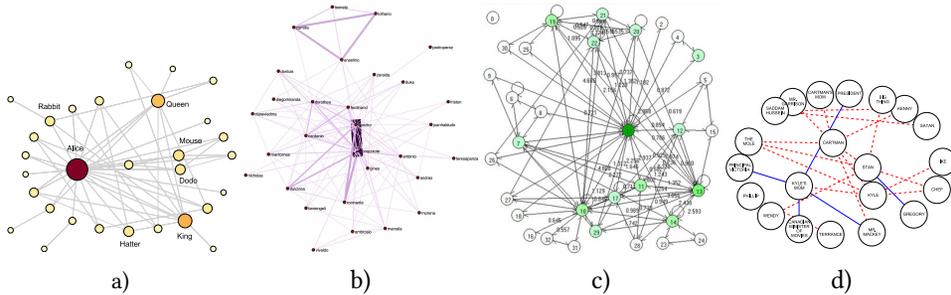


Fig. 7. Examples of the different types of edges found in the literature: a) *simple* character network of *Alice in Wonderland* [197]; b) *weighted* character network of *Don Quixote* [215]; c) *directed* character network of the movie *Avatar* [118]; d) *signed* character network of the movie *South Park: Bigger, Longer & Uncut* [129].

Interestingly, certain authors decide to ignore information that could be modeled as edge directions or weights, even if they have it at their disposal. This can be due to some methodological priorities, for instance, Yose *et al.* state that using directed edges to model unilateral interactions would result in a loss of statistical power when testing certain assumptions on the resulting network. The choice can also be caused by the will to adopt a simple approach. For example, Yose *et al.* justify their use of unweighted networks by the fact they are interested only in the presence of a relation between two characters, and not by its intensity [277].

The most important methodological issue one has to solve when extracting edges is to decide how to handle time. It is possible to consider the complete set of interactions occurring between two characters of interest over the narrative, and create a single edge summarizing all of them: this results in *static* networks (Section 5.1). However, this complete temporal integration leads to some information loss. Certain authors propose methods aiming at solving this issue, and produce *dynamic* networks (Section 5.2) instead.

5.1 Static Networks

In most of the approaches presented in the literature, the extracted character network is *static*, i.e. it represents the interactions between the characters for the period of time corresponding to the whole narrative. As interactions can occur anywhere in the storyline, obtaining such a network requires some form of temporal integration. On the one hand, the most widespread approach is to consider the set of successive interactions between two characters and derive an edge by applying a simple mathematical function, e.g. by counting them (Section 5.1.1). On the other hand, certain authors propose more advanced methods based on the explicit modeling of inter-character relationships (Section 5.1.2).

5.1.1 Count-Based Approaches. The most widespread methods to derive an edge from a series of interactions are relatively simple, relying on the existence or number of such interactions. Certain authors then discard some of the resulting edges, which they deem unreliable. Finally, a different extraction approach consists in extracting networks whose vertices represent groups of characters, and edges correspond to various forms of overlap between them.

least three [208, 214] or five [112] co-occurrences to keep an edge, whereas Stiller & Hudson just remove isolates [231]. Park *et al.* use a normalized weight, for which they empirically define a minimal threshold [189]. Sudhahar & Cristianini [236] do not deal with a single narrative, but with a *collection* of related stories, at once. They filter relationships depending on the number of documents in which they appear. The same method could be applied to subdivisions of a single fiction work.

Amancio [10] assesses the statistical relevance of a relationship through a null model. He works on co-occurrences, so his approach consists in comparing the edge weight (number of co-occurrences between the concerned characters) with the same weight computed over a randomly generated text. In [221], the authors transform their weights so that they represent distance in place of intensity. They then extract the minimum weighted spanning tree of their co-occurrence network, and use it in place of the original network, therefore discarding the rest of the edges.

Networks of character groups. Certain authors aggregate interactions, and more precisely co-occurrences, not only temporally, but also over *groups of characters*. Tsai *et al.* [252] extract a direct network in which a vertex corresponds to a group of characters co-occurring during the same scene, and possibly several ones. An edge going from one vertex to the other represents an inclusion between the character sets they represent. As described in Section 6.2.6, this type of network is designed specifically to summarize the plot. Grener *et al.* [99] proceed similarly, but an edge represents any non-empty intersection between two groups (instead of strict inclusion). Hettinger *et al.* [105] model temporal proximity rather than group similarity, by connecting two groups appearing successively in the plot.

5.1.2 Model-Based Approaches. Certain authors propose specific models to represent the relationships between two characters, taking advantage of the sequence of interactions between them, but also possibly of additional information such as the context of these interactions in the narrative. All these models rely either on co-occurrence- or conversation-related interactions.

Co-occurrences. Elson *et al.* [71] describe characters by vectors constituted of their numbers of occurrences during each unit of the considered narrative. They compute the correlation matrix over this representation of the character set, and use it as a weighted adjacency matrix to obtain what could be called an occurrence similarity network. Several authors propose to compare the context in which the characters occur in terms of topics. Kokkinakis & Malm [127] do so in Swedish novels: they characterize contexts through bags-of-words, and detect similar contexts using cluster analysis. They connect characters occurring in similar context, but also combine this information with other features, such as affiliations (see Section 4.5). Hutchinson *et al.* [110] first perform a Latent Semantic Analysis (LSA) to identify the lexical context of character occurrences in novels, followed by a Singular Value Decomposition (SVD) for dimension reduction. In the resulting matrix, each character is represented by a real vector. They then use the cosine similarity to compute the weight associated to each pair of characters. The result is a topic similarity network. Nan *et al.* [177] proceed similarly with TV series, by combining visual (objects surrounding the characters at the time of the occurrences) and textual information (subtitles).

Ding & Yilmaz [63] assume the existence of two conflicting groups of characters in the plot. As explained in Section 4.1.2, at the interaction detection step they leverage the narrative to compute a score representing the polarity of each scene. They then propose a Bayesian model taking into account the sequence of scene polarity scores and each character's group membership. It allows estimating the overall signed weight representing the polarity of the relationship between any two characters for the whole narrative, resulting in a signed weighted network. In [229], Srivastava *et al.* define a supervised approach to estimate the polarity of the relationships. They first extract

an unsigned graph based on co-occurrences, affiliations, and other interactions. They then train a classifier to predict the polarity of the relations, based on a number of textual features describing the context of the interactions, as well as structural features related to the notion of *structural balance* (categories of signed triads, cf. Section 6.1.5).

Conversations. Celiyilmaz *et al.* [45] focus on the content of the utterances. They propose a Bayesian model describing each character in terms of the topic about which he talks. From this, they extract build a topical similarity network: a fully connected graph in which the weight of the edge connecting two characters reflects how much they talk about the same things. According to the authors, this allows identifying hidden relationships between characters.

In [129], Krishnan & Eisenstein, take advantage of *how* two characters address each other over all their conversations, to estimate the nature of their relationship: *informal* vs. *formal*. The resulting network is therefore signed, positive and negative edges representing informal and formal relationships, respectively. In addition to linguistic content, their probabilistic model is able to take into account the structure of the graph, for example by assuming it is consistent with structural balance.

5.2 Dynamic Networks

Static networks present an obvious limitation, as identified by numerous authors in the context of character networks [11, 35, 184, 208, 247]: they result in a significant information loss, as they completely hide the chronology of the interactions. Yet, the order in which events occur is crucial to characterize a story, and for the writer it is generally at the core of the writing process. Moreover, the relationship between two characters is likely to evolve with the plot. This can lead to poor performance when solving certain problems based on a static network, or even completely prevent any resolution. Even from the descriptive point of view, Prado *et al.* [197] show empirically that the most central characters detected in static vs. dynamic networks can differ dramatically.

Certain authors try to cope for this by allowing multiple edges between vertices. For instance, in their signed network, Yose *et al.* [277] can have both a positive edge and a negative one connecting the same pair of characters, in order to model a relation evolving over time. Although simple, this is a very imperfect and incomplete solution. Another possibility is to extract a dynamic network. It takes the form of a sequence of character graphs called *times slices*, each one corresponding to a certain sub-period of the story. Like a static graph, a time slice is obtained through temporal integration, as it represents all the interactions occurring over a period of time. But this integration is performed at a much smaller time scale, which allows limiting the information loss. We distinguish two types of approaches: those using a fixed-size window to obtain these time slices (Section 5.2.1) vs. the others (Section 5.2.2).

5.2.1 Temporal Window. The fixed-size temporal window approach is clearly the most widespread. The notion of window depends on both the nature of the considered narrative and the utilization of the extracted network. The literature exhibits a variety of window sizes. For novels, most authors use one chapter [3, 11, 98, 197], as in Figure 9. Grener *et al.* [99] use a 1,000-word window, whereas Seo *et al.* [222] arbitrarily split the novel in ten equal-sized pieces. In theater scripts, Kwon & Shim [135] use a one-act window. For movies, authors use one scene [64], or automatically detected segments roughly corresponding to scenes [144]. For TV series, authors use one scene [35], one episode [21, 151] or one season [21]. Furthermore, certain authors use an overlap between consecutive windows, in order to preserve continuity as much as possible, e.g. for novels: fifteen pages with a fourteen-page overlap [208], 100 sentences with a ten-sentence overlap [184]. Weights can be processed for each slice, similarly to what is done for static networks (cf. Section 5.1.1).

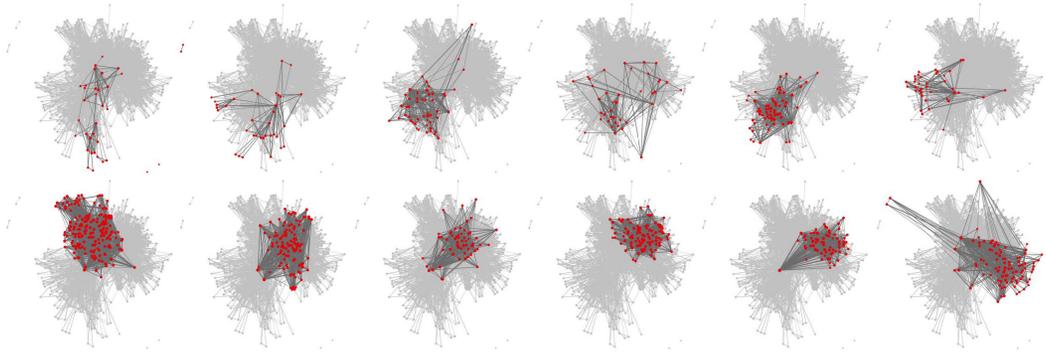


Fig. 9. *Dynamic* version of the character network of Rousseau's *Les confessions* [206], whose *static* version is presented in Figure 8a. Each time slice corresponds to a specific chapter.

Certain authors extract a dynamic but *cumulative* network [167, 175, 247, 264]. This means that, to obtain the graph representing a given moment in the narrative, they do not use only the corresponding time slice, but rather the portion of the narrative going from the very start to the considered moment, as illustrated by Figure 10. Other than that, the principle is the same as before: they perform a temporal integration over this period, and do so for every moment in the narrative to get a series of graphs constituting their dynamic network. As before, some authors consider unweighted graphs [264], others use the numbers of interactions [167, 247, 272] or the total interaction scores [245] as edge weights. Nalisnick & Braid [175] proceed similarly, but as their interaction scores are signed, the weights of the edges do not necessarily increase with time. On the contrary, one of the benefits of their approach is that a moment in the plot corresponding to a sign inversion for a given relationship is supposed to be narratively important for the concerned characters.

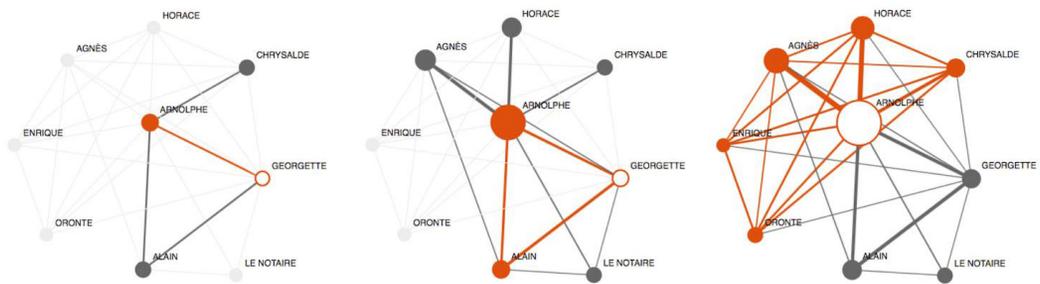


Fig. 10. *Dynamic cumulative* character network of Molière's *L'école des femmes* [272]. By construction, the size of the network increases over time.

Generally speaking, using a fixed-size window to discretize a time series requires knowing which window size to use. This is no trivial task, and it was shown that this parameter can have a very strong effect on the extracted network, and consequently on the process conducted on this network [56]. With character graphs in particular, on the one hand too large windows will miss many details, as relationships in works of fiction can evolve quickly, and on the other hand too small windows will result in very unstable networks, which may hide relevant information [184]. Bost *et al.* [35] have shown this experimentally, and also that a window too small can lead to mistaking

irrelevant events as important ones. Finally, a window too large may also cause problems when later studying the network: when the time slices differ much in terms of involved characters, it makes little sense to compare them through topological measures, as these are generally defined in a relative way [197]. Grayson *et al.* [98] propose to increase the window size until the graph density reaches a plateau, but this can produce very dense networks, likely to be uninformative.

5.2.2 Other Methods. In addition to the difficulties of estimating the best window size, there is no reason to suppose that this size should even be fixed: the tempo of the narrative is likely to change in any way, e.g. accelerating during action scenes and slowing down during emotional ones. A very few authors try to take this point into account.

Instead of considering fixed-size time slices, Mutton [173] uses an event-based approach to extract the sequence of graphs constituting his dynamic network. Each modification (creation or deletion of an edge, revision of a weight) results in the addition of a new graph in the sequence, representing the state of the character relationships after the modification. In some sense, this is the smaller possible time slice. Like some methods from Section 5.2.1, he adopts a cumulative approach. However, he also includes a temporal decay mechanism, which decreases the weights of an edge depending on the last time it was updated.

Another important point is that the narrative is linear, but not necessarily the story, or even the plot³. The narrative does not reflect the full state of all inter-character interactions at all times: it focuses on certain characters, presents under a sequential form some events actually occurring simultaneously, shows past events as flashbacks. A character can be absent from the narrative for a while, but still be active under the scene, performing actions that will be revealed later to the audience. Some authors consider it is necessary to model this hidden evolution to represent fully the character graph. Elsner [69] do so by applying a linear interpolation to determine the weight of a relation at a given point in the narrative for which this relation does not appear explicitly.

In [35, 36], Bost *et al.* propose *narrative smoothing*, an interpolation method allowing both to solve this problem and to mitigate the fixed-size window issue. Starting from the list of verbal interactions occurring at each scene, and weighted depending on their duration, they apply the following transformation to get their smoothed dynamic network. For a given pair of characters, if a verbal interaction occurs at the considered scene, the smoothed weight is the raw weight (duration). Otherwise, they take into account the last and next interactions of the considered characters: the raw weight is decreased depending on how much the characters interact with *others* in the meantime. Note that this decrease does not depend on the amount of time between two interactions of the considered characters, in order to avoid the fixed-size problem. This method allows estimating the weight of the relationship between any two characters at any moment of the narrative, even if there are not interacting at this time.

6 ANALYSIS AND APPLICATIONS

In this section, we describe how character networks can be used. We first review standard tools widely used to describe general networks, and discuss their interpretation in the specific case of character networks (Section 6.1). Then, we turn to real-world applications, and review the high-level problems solved thanks to character networks in the literature (Section 6.2).

6.1 Descriptive Tools

The literature contains a number of tools designed to characterize and study complex networks in general. Technically, most of them can be applied to character networks. However, their output is not necessarily relevant to this specific type of network, and when it is, some of them are used to

³The notions of plot (*what* is told), narrative (*how* it is told), and story (*perceived*) are defined in the introduction.

objectively describe these networks without providing any specific interpretation (e.g. centrality measures for automatic classification [264] or motifs for model fitting [29]). In this section, we review the topological measures used in the literature to describe character networks, focusing only on those for which authors propose a character network-related interpretation. We group them depending on the general approach they rely upon to describe the network.

6.1.1 Graph Density and Edge Weight. The most basic tools rely on counting the edges constituting the graph, or studying their weights (in the case of weighted graphs).

Graph Density. The density is the proportion of existing edges in the network, i.e. the ratio of its observed to possible edges. It is frequently used in the literature (together with other measures such as the number of vertices or edges) to compare networks extracted from several works, e.g. [71, 108, 115, 207]. A generalization of the density exists for weighted networks [98].

Certain authors directly and explicitly tie the density to some aspects of the narrative. Rochat & Triclot [209] use it to identify *behind closed doors* stories, which they associate to a high density (in addition to a structure devoid of any real center). Stiller *et al.* [231, 232] study the evolution of the density as a function of the number of characters in Shakespeare’s plays. They show that it decreases, which they tie to the will of the writer not to involve too many characters in his subplots, in order to keep the story understandable. Together with a moderate transitivity, Sack [214] interprets low density as a marker of social dissociation, i.e. presence of largely disconnected social groups.

It is worth noting that, despite the apparently plot-related interpretation of density, this measure is also used to characterize stylistic aspects of writing. For instance, Voloshinov & Gozhanskaya [261] use the density to discriminate plays by different playwrights.

Edge Weight. By construction, a weight (the numerical value associated to an edge in weighted networks) represents the intensity of the relation modeled by the concerned edge. In the literature, it is considered from a comparative perspective, according to two modes: between characters and over time.

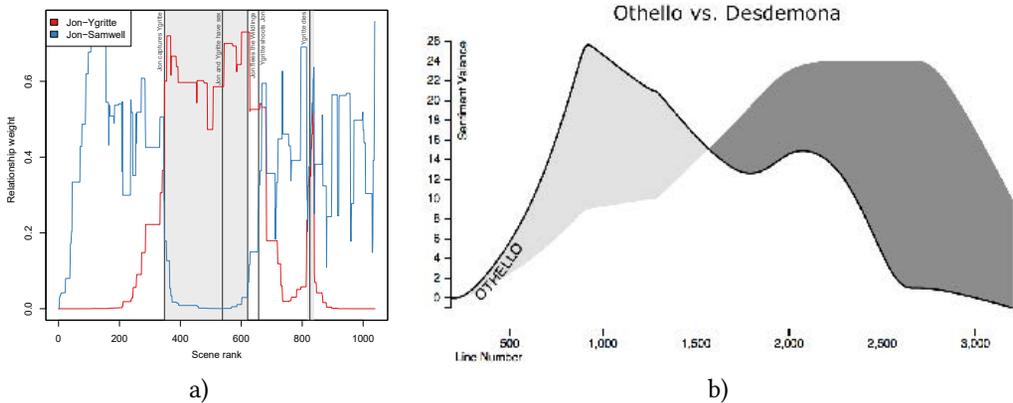


Fig. 11. a) Evolution of the link weights for three important characters in the *Game of Thrones* TV series [36]; b) Evolution of the signed weight of a relationship of interest in *Othello* [175].

Certain authors distinguish between weak and strong edges, depending on their values. Sack [214] observes that a low proportion of the latter indicates infrequent and superficial interactions (as their weights correspond to interaction counts). These are often taking place between the protagonist and punctual minor characters, in order to provide variety, color, and some background to the

story [216]. On the contrary, a large proportion of strong edges is indicative of complex and durable relationships, developed over a long period of time [216]. Suen *et al.* focus only on the strongest edges: they propose to compare the top two weights of the network, in order to assess how much the plot focuses on the most important relationships relative to the others. It is also possible to consider the whole distribution, as Gleiser does for comics [92]: he identifies a power law, and interprets it as revealing the presence of a few relationships largely dominating the narrative space (e.g. Spider Man and his wife).

Instead of comparing the weights of the edges connecting different characters, some authors focus on a relationship and study how its weight evolves through time (cf. Figure 11). This is what Bost *et al.* [35, 36] do for TV series, interpreting significant changes as important event occurring in the narrative. Nalisnick & Braid [175] adopt the same principle, with the difference that their weights are signed. They interpret a sign change as an event of importance. For instance, in Shakespeare's *Hamlet*, the moment where the sign of the relationship between the eponymous character and his mother switches from negative to positive corresponds to a crucial revelation regarding the death of their father/husband.

6.1.2 Degree and Strength. The *degree* of a vertex is the number of edges attached to it. It is sometimes normalized to get a value in $[0; 1]$, as in [231]. If the considered network is directed, one can distinguish the *incoming* from the *outgoing* degrees, which correspond to the number of incoming and outgoing edges attached to the considered vertex, respectively. The *strength* is the weighted generalization of the degree: the sum of weights over the edges attached to the vertex. A high strength can be due to few edges with high weight, but also to many ones with small weight, so both measures are not necessarily correlated.

Hubs. A *hub* is a vertex with a high degree. In a character network, it is interpreted as a pivotal character, one that interacts with many others. Depending on how the network is extracted, this large number of interactions can be spread over time, or concentrated in a few events involving many characters. When considering strength, the idea is the same as for the degree, but considering the intensity (or whatever property conveyed by the weights) of the exchanges. Many authors use these measures to identify primary characters in the narrative. For instance, Hamlet has the highest degree in Shakespeare's *Hamlet* [170], and the hero always has the highest degree in the collection of stories and tales analyzed in [236]. When studying Bengali Literature, Muhuri *et al.* [172] observe that both protagonists and antagonists have a significantly higher strength than minor characters. If the narrator is a character, she is probably a hub, as she will tell the story from her perspective, and is therefore likely to intervene more [206]. Of course, it is possible for a character network to contain no hub at all, e.g. in the case of a complete network (all characters are interconnected), or to contain several hubs. Agarwal *et al.* [3] propose a method to identify the narrator (see Section 6.2.3).

In directed networks, the incoming and outgoing degrees or strengths are often seen as popularity and sociability markers [172], respectively. It is also possible to interpret them in terms of dependency and force for change, as an incoming hub is someone towards which other characters turn to in case of need, and an outgoing one denotes a will to act on the environment. For instance, in their directed mythological network, Choi *et al.* [53] identify Zeus as an incoming hub (*in charge* for the whole world) and Heracles as an outgoing one (the *cause* of many great events). Sudhahar & Cristianini [236] make a similar observation for stories and tales, as heroes generally have a higher outgoing strength (they tend to be the agent of many actions). However, they also notice that strength (or degree) alone may not be enough to identify primary characters, since other authoritative characters (e.g. God, the King) can be as much (or even more) of a hub.

When a dynamic network is extracted, certain authors study the evolution of the degree to study the importance of the characters. For example, Rochat [206] shows that in Rousseau's *Les*

confessions, the transition from the first part of the novel to its second part corresponds to some sort of passing of power between two important characters. Agarwal *et al.* observe that ALICE has not always the highest degree in *Alice in Wonderland* [3]. Min & Park show that in *Les miserables*, the focus alternates between VALJEAN and MARIUS. Bost *et al.* [35, 36] leverage degree evolution to detect important events, such as the death of a character.

Distribution. Alternatively, instead of focusing on individual vertices, it is standard in the domain of complex network analysis to study the *distribution* of the degree and/or strength. Authors generally observe either an exponential [154] or a power law distributed degree, for undirected [8, 135, 189, 228] as well as directed [53, 136] networks (cf. Figure 12a and 12b). The power law is frequently found in other (non-fictional) social networks. In character networks, it tends to appear when the narrative is dominated by one or a few characters. A network displaying such a degree distribution is called *scale-free* [17], and this property has been quite widely observed in many types of real-world networks. This supposed ubiquity has been criticized lately though [42, 138], as properly determining if a sample follows a power law is difficult in practice.

With a power law distribution, the average is not a very informative statistic to describe the degree, as it is not a characteristic value. However, authors traditionally like to indicate it [108, 137, 154], sometimes with the minimal and maximal degree values [8, 277], standard deviation [214], variance [238], or median [242]. In [71, 115], the average degree and strength are two among several features used to compare networks extracted from distinct works.

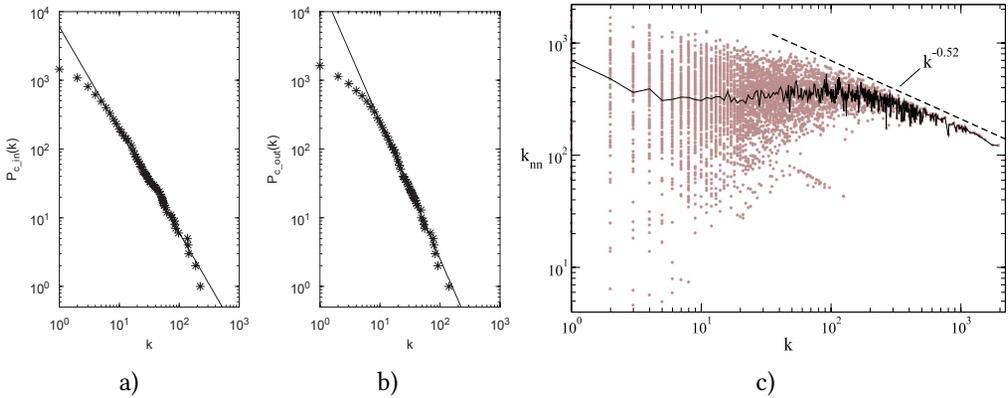


Fig. 12. a) Complementary cumulative in-degree distribution for the *Greek and Roman mythology* network [53]; b) Complementary cumulative out-degree distribution for the same network [53]; c) Mean degree of the neighbors of a vertex as a function of its degree, for the *Marvel* network [92].

Other statistics are designed to summarize the degree distribution, especially its heterogeneity. Suen *et al.* [238] propose the *Single character centrality*, which is not unlike Freeman's *Centralization* [83]. It is based on the comparison of the network highest and second highest degree/strength, normalized by its total degree/strength. By comparison, the centralization is defined generically, for any vertex centrality measure, and compares how central the most central vertex of the network is, relatively to the rest of the vertices. The Single character centrality allows quantifying how much the top character dominates the others. In order to assess the spread of the distribution, they also propose to use the entropy, and the degree/strength mean and variance for the network restricted to its top 10 characters. The same statistics could, of course, be applied to any other centrality

measure. Suen *et al.* use all of these as features to train a classifier in distinguishing predefined classes of plots (cf. Section 6.2.4).

Assortativity. The *degree assortativity* [178] measures how correlated the degrees of connected vertices are. Real-world social networks are usually assortative, i.e. their vertices tend to be connected to other vertices of similar degree [180], but this is not the case for many other types of real-world networks (e.g. technological or biological networks). Rochat uses assortativity [206] to characterize the overall relationship between primary and minor characters. In the case of Rousseau's *Les confessions*, he finds a slightly disassortative network, indicating that the edges tend to connect a small set of primary characters present most of the narrative, to a large set of minor characters merely passing by. Moreover, it also shows that major (resp. minor) characters do not tend to connect to each other.

Similarly, Mac Carron & Kenna [154] observe that certain fictional character networks tend to be disassortative. They make the same observation for epic narratives, yet these are not necessarily (completely) fictional. They assume this is due to the conflictual nature of their plots: a lot of characters are introduced just to be killed by heroes, resulting in many small degree vertices connected to large degree ones. Indeed, when ignoring this type of interactions, the networks are assortative. They use this observation to support the assumption that some aspects of these tales are real, whereas others are fictional.

6.1.3 Shortest Path and Distance. The vertex-to-vertex distance, a.k.a. the geodesic distance, or *distance* for short, is very popular to describe networks in general. It corresponds to the length of the shortest path between two vertices. Several popular topological measures are based on the notions of shortest path or distance.

Distance-Based Statistics. As a part of the *small world* property [263], it is often observed that the distance averaged over all pairs of vertices (or *average distance*) is relatively small in real-world networks, and increases only slowly (logarithmically, to be precise) with the number of vertices, in growing networks. In such small world networks, the *diameter*, which corresponds to the maximal distance in the network, has roughly the same behavior and is therefore alternatively studied. Most authors examining the average distance or diameter in character networks find small values, e.g. [8, 29, 53, 154].

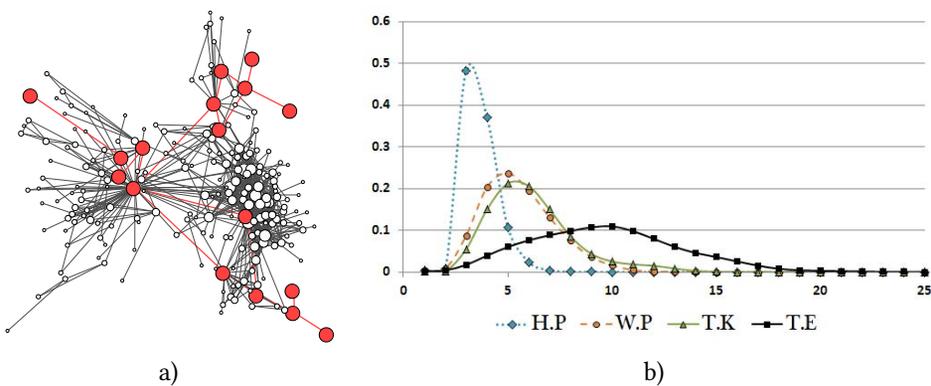


Fig. 13. a) All diameter paths (in red) of Rousseau's *Les confessions* (cf. Figure 8a) [206]; b) Distribution of character-to-character distance in 4 novels (*War and Peace*, *The Three Kingdoms*, *Harry Potter and the Philosopher's Stone*, *The Earth*) [221].

In the context of character networks, both the *average distance* and *diameter* (maximal distance) are used to determine the compactness of the story (cf. Figure 13). A large diameter, with respect to the network size, means that at least certain characters are narratively well separated, whereas a large average distance denotes a more general trend. Rochat [206] interprets a small diameter as a strong overlap between the characters, resulting in a narrative proximity between characters, even those with different profiles, roles, or appearing at different moments. He also studies the nature of the different shortest paths whose length corresponds to the diameter, and observes that in this case, they connect minor characters by passing through primary ones. Using Cervantes’ *Don Quixote* as an example, Sack [216] observes that a large diameter is indicative of an episodic narrative, i.e. one where the protagonist successively meets independent groups of characters. This large diameter then goes with a high transitivity (Section 6.1.5), reflecting the locally cliquish structure of the network. Incidentally, such a high transitivity is also a prerequisite to Watt & Strogatz’s *small world* property [263].

Betweenness Centrality. The *betweenness* of a vertex [82] is related to the number of shortest paths going through this vertex. In a character network (cf. Table 1), it allows measuring how much a character acts as a broker or narrative bridge, i.e. connects separate parts of the plot [26, 206]. Characters with low betweenness are considered as narratively dependent on more central ones [206]. However, Bolioli *et al.* remark that the most central characters in terms of betweenness are not necessarily those appearing as the most important to the reader, as this centrality highlights their bridging role [26]. Muhuri *et al.* also remark that a high betweenness is associated to a role of messenger, generally towards supporting characters [172]. In the collections of stories and tales studied by Sudhahar & Cristianini, heroes have the highest betweenness centrality [236]. The edge-betweenness measure is an edge-based variant of the measure⁴, and allows highlighting important narrative transitions [206].

The whole network can be described through betweenness-based statistics, for instance by averaging it over all vertices, or by computing the betweenness-based centralization [83]. In [207], Rochat computes this centralization for all 20 novels of Zola’s *Les Rougon-Macquart* series of novels, in order to determine which ones are centered on a single protagonist, or revolve around several ones (a type of plot called *polyfocalisation*). Suen *et al.* [238] proceed like they do for the degree, i.e. using the maximum value, the difference between the 2 top vertices, and the entropy to summarize the betweenness distribution.

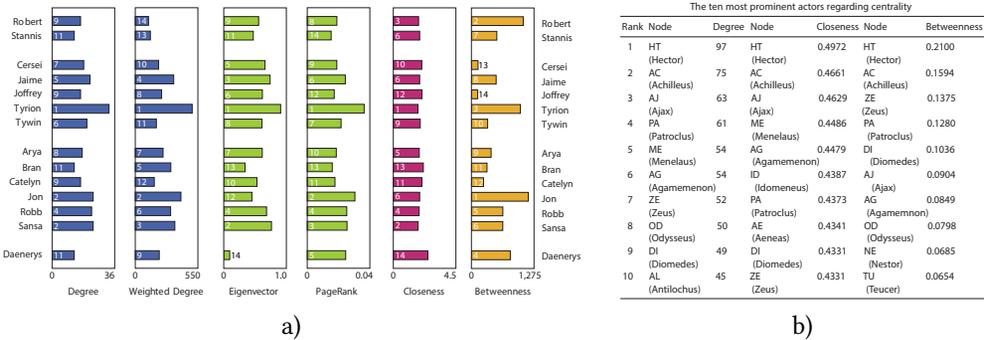


Table 1. a) Various centrality measures computed for the network of *Game of Thrones* [23]; b) and for the network of *The Illiad* [137].

⁴Edge-betweenness: number of shortest paths going trough some edge of interest.

Closeness Centrality. The *closeness* of a vertex [20] is the reciprocal of the total (or sometimes average) distance between this vertex and the rest of the network. In a character network (cf. Table 1), it can be considered as measuring how close a character is from all the others, narratively speaking [206], i.e. in terms of their respective character-spaces.

Alberich *et al.* [8] use the closeness to identify the center of their *Marvel universe* network (which turns out to be Captain America). Rochat notes that the variance of this measure is usually low, which makes it harder to discriminate vertices. He prefers to use the *harmonic* closeness (based on the harmonic mean) [25] instead, which additionally allows handling disconnected networks. Rochat uses it to identify important minor characters that, although not proper protagonists, are still involved in non-anecdotic branches of the story.

Some authors alternatively use a related measure, which has a similar interpretation. The *eccentricity* of a vertex is the distance to its farther vertex [101]. It can be used to determine the center of the network (vertex of lowest eccentricity), or determine how many centers it contains. For instance, in [206] Rochat identifies MME DE WARENS as the only center in Rousseau's autobiographic novel *Les Confessions*, as she is, by far, the least eccentric vertex (in the story, she is his mentor). In [53], Choi *et al.* consider both the maximal (eccentricity) and minimal distance to a few vertices of interest, to further describe their position.

6.1.4 Connectivity. It is possible to characterize a network by studying its *connectivity*, either directly or through derived measures. A graph is said to be *connected* if there is a path connecting each pair of its vertices, and disconnected otherwise. A disconnected graph is consequently made of several separated *components* (which are themselves maximally connected subgraphs). Real-world networks generally possess a so-called *giant component*, which contains almost all the vertices.

Components. In the context of character networks, the meaning of the connectivity depends on several aspects of the extraction process: raw data, type of interactions modeled by the edges, and whether the network is static or dynamic. Suppose the network represents mentions or affiliations: it is very possible to have separated components, for example because the plot contains two opposed families. But if the edges model conversation or other interactions, authors generally observe a giant component [155]: indeed, for such edges, a disconnected network would mean there are completely separated subplots.

This is unlikely, but possible though: for instance, Mac Carron & Kenna [154] observe that the epic narrative *Beowulf* contains two subplots taking place in the past, and completely disconnected from the main plot. They appear as two separate components, whereas the main plot is a giant component. It is also possible that the raw material is not a single work, but rather a collection of works. It is the case, for instance, with the comic books composing the Marvel universe, which leads to a disconnected network dominated by a giant component [8]. It is also likely to see such disconnections if the network is dynamic: one can observe several separated components at some point, representing several subplots likely to merge later in the story. In the presence of a giant component, some authors choose to ignore isolates and small components [136, 137], which are considered as negligible, as is generally done more generally in the literature when studying other types of complex networks.

Articulation Points. The importance of a vertex (or a group of vertices) in a graph can be studied by assessing how its removal affects the graph connectivity. This is related to the notion of *articulation point*: a vertex whose removal makes the graph disconnected. Several authors adopt this approach: Moretti [170] to study Shakespeare's *Hamlet*, Falk [75] for Edgeworth's *The absentee*, and Mac Carron & Kenna [154, 155] for mythological works. Moretti removes various combinations of the most central vertices in terms of degree or closeness, in order to study which ones split the

network. He notes that a protagonist is significant mainly because of the stability he brings to the network. Moreover, he shows through a series of examples that this quality is not completely reflected by these centrality measures, as the removal of similarly central characters differently affects the network. He favors transitivity to explain how the removal of these characters leads to different results, as some of them are part of a dense, highly transitive part of the network, whereas others are not. Mac Carron & Kenna observe that removing the few top central vertices in terms of betweenness [154] or degree [155] is enough to disconnect their mythological networks, whereas they are robust to random deletion. The authors use this to highlight the importance of a few main characters: the story just falls apart without them.

In [128], Koschützki *et al.* propose to measure the importance of a vertex based on how its removal affects a graph statistic of interest. For this purpose, one computes the statistic for the whole network, then for the same network without the vertex of interest. The difference between the two resulting values is called *Vitality*. Rochat adopts this approach [206] when studying Rousseau's *Les confessions*, using centralization as the graph statistic. Rochat applies it to the degree, closeness, betweenness and Eigenvector centralities, in order to study the vitality of the vertices. He also computes the vitality using the number of components as the graph statistic, which allows him to detect articulation points.

6.1.5 Triads. Several standard topological measures are based on the notion of triads, i.e. sub-graphs constituted of three vertices. We focus here on *transitivity* and *structural balance*.

Transitivity. The *transitivity*, a.k.a. *clustering coefficient* is a measure defined at two different levels. At the *global* one [152], it is the proportion of existing triangles (i.e. closed triads) relatively to the possible ones (open triads). More formally, it is the probability, when randomly selecting one vertex and two of its neighbors, that these neighbors are themselves connected. Some authors adopt a relatively close approach by counting the number of triangles in the graph [71], as illustrated by Figure 14. At the *local* level [263], the transitivity is defined relatively to a vertex of interest: it is the proportion of its neighbors forming connected pairs. Put differently, it is the probability, when randomly picking two neighbors, that there is an edge between them. The *small world* property mentioned before (Section 6.1.3) requires that, in addition to a small average distance, the considered network exhibits a high transitivity when compared to an equivalent random (Erdős-Rényi – ER) network [263]. A number of authors observe such a high transitivity in character networks [8, 11, 29, 154, 232]. A weighted generalization of the transitivity exists, which is occasionally used, e.g. [238].

When studying Shakespeare's plays, Stiller *et al.* [232] associate a high *global* transitivity to the presence of separated social groups always appearing on stage together, thereby forming cliques. In a subsequent work [231], Stiller & Hudson show that certain characters have a low *local* transitivity though, because they appear in several such subgroups (so, many of their neighbors are not connected). These characters, which they call *keystones*, are very important in the plot structure, as they provide some continuity by connecting subplots. Other authors observe a strong inverse correlation between degree and transitivity in various fiction works [53, 92, 154]: hubs tend to be keystones, whereas minor characters are well embedded. In [172], Muhuri *et al.* observe that the supporting characters of antagonists are more tightly knit than those of protagonists, leading to higher transitivity for the former.

Certain authors use the local transitivity to determine the hierarchical nature of their network: *Marvel universe* [92], mythological works [53, 154, 155, 168]. They observe that the transitivity considered as a function of the degree takes the form of a power law. This means that vertices with a small degree have a much higher transitivity, and that hubs have a very small one. Narratively speaking, minor characters tend to form small dense clusters, resulting in high transitivity but low

degree, whereas major characters act as bridges and connect these clusters, which results in high degree (many clusters to connect) but low transitivity (these clusters are mutually disconnected). In the case of Choi *et al.*'s Greek/Roman mythological network [53], the authors assume that this is mainly due to two reasons. The first is the largely genealogical nature of the relationships between characters, resulting in a tree-like structure. The second is the presence of classes of characters (gods, titans, monsters, etc.) shaping the narrative of the myths.

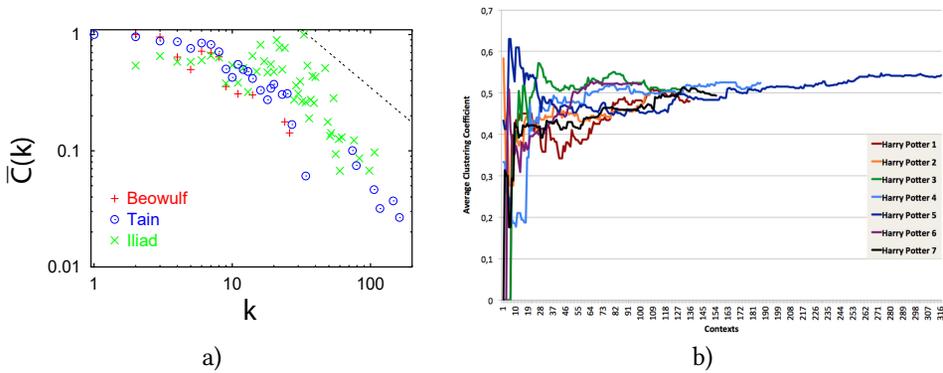


Fig. 14. a) Average local transitivity as a function of the degree, in 3 mythological narratives (*Beowulf*, *The Iliad*, the *Táin Bó Cúailnge*) [154]; b) Average local transitivity as a function of time, in the cumulative networks of each *Harry Potter* novel considered separately [264].

Structural Balance. Like the transitivity, the notion of *structural balance* is related to triangles (closed triads), but in *signed* networks this time. There are several variants of the structural balance concept, the simplest being: a triangle is structurally balanced if it contains exactly zero or two negative edges [44]. Consider for instance a social network, in which positive and negative represent friendly and hostile relationships, respectively. In a balanced triangle, all three persons can be friends, or two can be friends while being both hostile towards the third. The case where all three persons are mutually hostile to each other is considered as imbalanced, as two of them are expected to eventually unite against the third, resulting in the previous situation. Similarly, when one person is friend with two mutually hostile persons, he might start a fight with one of them, or both of them might become friends, also resulting in a balanced situation. Real-world social networks were initially expected to contain mainly balanced triangles, although this assumption has been put in question lately [66].

In the case of character networks, edge signs represent antagonistic relationships, generally friendly vs. hostile [63, 151, 157, 175]. Mac Carron *et al.* [154, 155, 157, 277] study the number of imbalanced triangles in their mythological networks, and the low proportion they find is comparable with what is generally observed in real-world social networks. Sack [214–216] studies the evolution of the network *balance*, a quantity related to the number of balanced triangles it contains, to identify several phases in the plot (progress towards stability, reversal, factions formation, resolution...). Their approach relies on the Aristotelian structure, and they assume that the character network is initially imbalanced, and that the plot drives it towards perfect balance (the story resolution).

However, Liu & Albergante [151] make an opposite observation when studying the *Game of Thrones* TV series, as they observe around 30% of imbalanced triangles. They assume that the writers voluntarily keep the plot in this imbalanced state in order to make it more interesting for the viewer. When comparing episodes, they also remark that increases in the number of edge changes during

an episode (be it the inversion of a sign or the addition/removal of an edge) are related to higher viewer figures. They assume that such changes reflect how much the plot progresses, and that an important evolution is likely to strongly engage the audience.

6.1.6 Modular Structure. A character network can generally be broken down into a collection of disconnected or loosely connected subnetworks, corresponding to its modular structure.

Community Structure. A network is said to possess a *community structure* when it can be partitioned into subgraphs called communities, such that many edges fall inside these groups and few between them [88]. Community detection is an important problem in the domain of complex network analysis, and a number of methods have been proposed for this purpose (cf. [81] for a detailed review).

In character networks, communities are related to parts of the narrative, as illustrated by Figure 15. They can be used to identify either individual subplots, or groups of subplots involving roughly the same characters [26, 136, 155, 181, 191]. For instance, in the epic narrative *Beowulf*, Mac Carron & Kenna [154] identify two subplots taking place in the past and appearing as two components, completely disconnected from the main plot. It is also likely to see such disconnections if the network is dynamic: one can observe several separated components at some point, representing several subplots likely to merge later in the story. But the different subplots are generally at least weakly connected, resulting in a community structure [81] instead of several distinct components.

Several authors observe that such communities gather characters based on spatial and/or temporal criteria. This is for instance the case in Rousseau's *Les confessions* [206], Martin's *Game of Thrones* [23], and certain Icelandic sagas [155]. Jayannavar *et al.* [115] observe that, in networks extracted from 19th century British novels, the number of communities increases while their average size decreases when the network gets larger in terms of characters. This would be related to writing constraints, as it is not easy to handle large groups of interacting characters without confusing the reader.

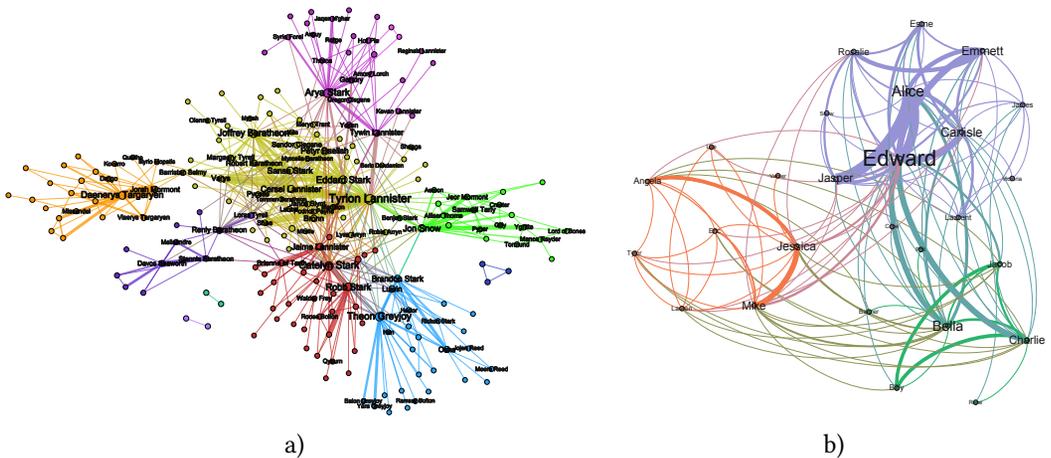


Fig. 15. a) Communities in the TV series *Game of Thrones* (seasons 1 & 2) [35]; b) Communities in the three novels of the series *Twilight* [29].

For Lee *et al.* [144], in the case of movies, each community is centered on a main character, accompanied by sidekicks and other types of minor characters. However, this is not the consensus, as several authors note that primary characters are likely to belong to several communities, because

they appear in a number of subplots [172, 191, 267]. In particular, Muhuri *et al.* [172] propose a custom measure called *Diversity*, in order to quantify how well connected a character is to the communities of the network. It is similar enough in principle to the *Participation Coefficient* [100] and *Diversity Score* [67] proposed to study other types of complex networks. They observe that main characters tend to be connected to multiple communities, unlike minor ones. Implicitly, all these authors identify the need to apply a community detection method able to identify *overlapping* communities, however none of them apply such a tool.

Correlation Clustering. When the network is signed, its modular structure is detected by solving the *Correlation clustering* problem [16], which consists in finding groups such that most positive and negative edges are located inside and between the groups, respectively. The result is interpreted in terms of antagonistic groups rather than subplots. Ding & Yilmaz [63] adopt this approach to identify opposed factions in so-called *adversarial* movies.

6.1.7 Miscellaneous Approaches. More minor methods are used in the literature to described character networks, which we summarize in this last section.

Spectral Measures. A family of vertex centrality measures is based on the spectrum of the graph adjacency matrix, or one of its related matrices (e.g. graph Laplacian), cf. Table 1a and Figure 16a. The *Eigenvector centrality* [27] can be considered as a generalization of the degree centrality, in which instead of counting the neighbors of a vertex, one also considers their respective centrality values. Rochat [206] notes that this measure tend to be higher in the center of dense communities, such as those corresponding to subplots revolving around multiple characters. By opposition, if the subplot is highly centralized around a single character (such as Mme de Warens in Rousseau's *Les confessions*), the Eigenvector centrality is lower.

In [63], Ding & Yilmaz study what they call *adversarial* plots, which oppose two groups of characters. After having first bisected their network to estimate these two groups (cf. Section 6.1.6), they use the Eigenvector centrality to identify the leaders of these communities. These are then assumed to be the hero and the villain of the story. In [197], Prado *et al.* compute the *Centralization* [83] based on the Eigenvector centrality. They consider it allows measuring how close a graph is from a star structure. They associate this star structure to single character-centered plots, such as biographies. They also use a dynamic generalization of the Eigenvector centrality to describe the trajectory of individual characters.

HITS is a generalization of the Eigenvector centrality to directed networks [125]. It includes an *Authority score* and a *Hub score*, depending on whether one considers incoming or outgoing edges. Certain authors use them them to identify the main characters in novels [3, 236].

k-Cores. Cliques and various relaxations of this concept can be used to describe a graph. The notion of *k-core* refers to a maximal induced subgraph whose all vertices have a degree of at least k [220]. It is obtained by iteratively removing isolates and leaves. Park *et al.* [188, 189] use it to study the structure of novels, and distinguish what they call the *kernel* from the more peripheral characters. They proceed similarly in a subsequent article [221], but using a filtered version of their co-occurrence network, taking the form of a minimum weighted spanning tree.

The *k-core number* (or *coreness*, or *degeneracy*) of a graph is the largest k for which this graph contains a k -core. Rochat [207] proposes to normalize it relatively to the size of the network, in order to measure the compactness of the group of main characters.

Cyclic Coefficient. The *cyclic coefficient* was defined by Kim & Kim to measure what they call the degree of circulation of a network [124]. It quantifies how cyclic vs. treelike a structure is. It is defined relatively to a given vertex. One first identifies the set of shortest cycles going through the

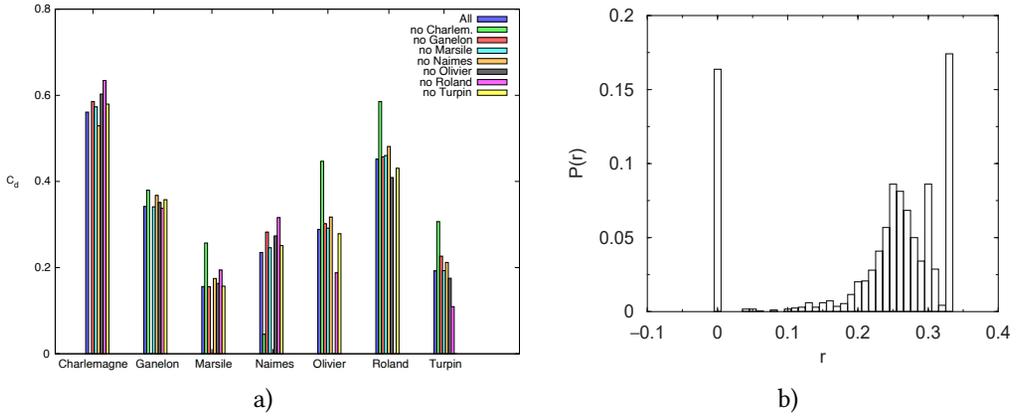


Fig. 16. a) Eigenvector centrality for the main characters of the song of heroic deeds *La Chanson de Roland* [197]; b) Distribution of local cyclic coefficient for the *Greek and Roman mythology* network [53].

vertex and each pair of its neighbors, then averages the reciprocals of their lengths. The resulting value ranges from zero (the graph is locally a tree) to $1/3$ (a clique). The authors propose to average this measure over V in order to describe the whole graph.

In [53], Choi *et al.* use the average cyclic coefficient to describe the network they extract from a biographical dictionary of Greek and Roman myths (cf. Figure 16b). This allows them to confirm their assumption that this network is treelike due to its mainly genealogical nature. They also study the distribution of the cyclic coefficient and identify two classes of characters: very peripheral characters embedded in a locally treelike neighborhood, vs. characters whose narrative is more or less tightly intertwined with others', and belonging to clique-like structures.

Vertex Attributes. Certain authors extract character networks possessing vertex attributes, i.e. each vertex is described by certain individual fields such as gender, age, or social group. This information is leveraged in different way when studying the network. The simplest is to assess the prevalence of the attribute values in the graph. When studying their science-fiction corpus in [209], Rochat & Triclot do so with the vertex attribute describing the main occupation of each character: scientist, politician, technician, religious figure, animal, etc. For instance, they observe that the most widespread occupation among the main characters is scientist.

If the network is large enough and sufficiently connected, it is also possible to study the induced subgraphs corresponding to sets of vertices possessing certain attribute values. On the *Illiad*, Kydros *et al.* [137] consider four categories of vertices (*Greeks*, *Trojans*, *Gods*, and *Others*). They study them separately, by computing a collection of standard topological measures on the induced subgraphs. For instance, they find out that the Trojan subgraph is much denser than its Greek counterpart, revealing a much higher cohesion of these characters in the story. They also study how these four groups of vertices are interconnected. As an example, the number of relationships between gods and mortals are similar for either side, denoting a carefully balanced interventionism.

Another approach is to compute the *assortativity* (or *homophily*) of the network for an attribute of interest. It is obtained by applying an association measure to the two series constituted by the attribute values of *connected* pairs of vertices [179]. In the case of numerical attributes, one generally uses Pearson's correlation coefficient, and for categorical ones it is Cohen's κ . Put differently, the assortativity measures whether the edges tend to appear between *similar* vertices (positive value), in terms of the considered attribute, or *dissimilar* ones (negative value). In the first case, the network

is said to be *assortative* (or *homophilic*), and in the latter, it is *disassortative* (or *heterophilic*). Yose *et al.* compute the assortativity of the network extracted from the Irish epic text *Cogadh Gaedhel re Gallaibh*, which describes a war between Irishmen and Vikings. They leverage this measure to determine whether hostile interactions tend to be *internal* (i.e. connect Irishmen), which would mean the text mainly describes a domestic conflict, or *external* (i.e. connect Irishmen and Vikings), which would mean both sides are opposed. Bossaert & Meidert [32] use a similar approach to check whether adolescent support is affected by school year, school house or gender in the *Harry Potter* series of novels.

6.2 Applications

In this section, we present a selection of methods and results aiming at solving a specific problem and relying on character networks for their purpose. We overlook the many articles from the field of Literary Studies, which focus on a single fiction work (or a few ones) and aim at describing it in great details [206, 208]: these are case studies, their goal is not to propose a generic tool that one could subsequently apply to other works. We also ignore articles only using character networks for visualization [99, 173, 259, 272], as they are somewhat limited, at least from this perspective.

6.2.1 Assessment of Literary Theories. Some articles aim at assessing the validity of literary theories. Those are often evaluated qualitatively and/or on a small number of fiction works [71]. Automated approaches can help at testing them in a more quantifiable way, and/or on larger corpora.

In [71], Elson *et al.* focus on two literary assumptions related to the size of the community surrounding the protagonist in 19th century British novels. The first one is that it depends on the amount of dialogue occurring: it would not be possible to show many conversations when there are many characters to consider. This is formally translated as the presence of a strong correlation between the amount of dialogue and the number of characters in the novel. The second assumption is that it depends on the social setting of the novel: there would be more verbal interactions in rural than urban communities. This is handled by manually determining the setting of the considered novels.

Elson *et al.* adopt conversational networks to model the novels. They use a number of features to assess both literary theories, some describing the network (numbers of 3- and 4-cliques, average degree, density) and others extracted from the detected verbal interactions (e.g. numbers of speaking and non-speaking characters, variance of the number of utterances by character). They apply their method to their corpus of sixty classic novels. They find a slightly positive correlation between the numbers of characters and utterances, and an even stronger one when counting only speaking characters: this invalidates the first literary theory. The same behavior is observed when considering the average degree and the density instead of the number of utterances, which the authors consider as a confirmation. Regarding the second theory, Elson *et al.* do not observe any significant difference in the numbers of characters (speaking or not), or any tested feature, relative to the type of setting (rural vs. urban), which invalidates the second theory. Instead, the most influential factor seems to be the narrator's perspective: characters are much loosely connected when the story is told in the first person than in the third.

Jayannavar *et al.* [115] later revisit these results, based on SINNET, the tool developed by Agarwal *et al.* [1, 3–5]. It allows identifying two kinds of interactions: unilateral vs. bilateral interactions (cf. Section 4.4 for more details), the former supposedly subsuming conversations. The authors argue that identifying a wider range of interactions might change the results. Moreover, they observe that Elson *et al.* misunderstood the second assumption: they confused the notions of literary *setting* (the real world environment) and *space* (how it is portrayed in the novel, be it accurate or not).

topological properties similar to real-world social networks. Stiller *et al.* [232] study a corpus of ten plays by Shakespeare in an effort to determine whether the success of this playwright relies on his ability to mimic some basic properties of real social networks in his writing. Alberich *et al.* [8] and Gleiser [92] assess how realistic cooperation between *Marvel* characters is, through the comparison of their co-occurrence network with real-world collaboration networks (e.g. IMDb, DBLP). In a series of articles studying myths, tales and other mythological stories, Mac Carron *et al.* [121, 122, 154, 155, 157, 158, 277, 278] assume that such narratives convey both actual and fictional information. They assess their historicity by determining where they stand between completely real and purely fictional social networks.

Play	C	D	T	Network	N	$\langle k \rangle$	ℓ	ℓ_{rand}	ℓ_{max}	C	C_{rand}	G_c	r
Hamlet	0.25	1.8	0.82	All	74	4.45	2.37	2.88	6	0.69	0.06	50 (67.5%)	-0.10
				Hostile	31	1.67	2.08	3.25	4	0	0.05	10 (32.2%)	-0.20
				Friendly	68	4.12	2.45	2.98	6	0.69	0.06	45 (66.1%)	-0.03
King Lear	0.39	1.76	0.76	All	404	6.10	2.76	3.32	7	0.82	0.02	398 (98.5%)	-0.33
				Hostile	144	2.33	2.93	5.88	7	0.17	0.02	131 (90.9%)	-0.36
				Friendly	385	5.67	2.84	3.43	7	0.84	0.01	350 (90.9%)	-0.32
A Midsummer Night's Dream	0.51	1.57	0.87	All	716	7.40	3.54	3.28	11	0.57	0.01	707 (98.7%)	-0.08
				Hostile	321	2.25	4.10	7.12	9	0	0.01	288 (89.4%)	-0.39
				Friendly	664	6.98	3.83	3.34	12	0.62	0.01	547 (82.3%)	0.10
Othello	0.50	1.55	0.72	All	67	3.49	2.83	3.36	7	0.68	0.05	43 (64.2%)	0.01
				Hostile	324	3.71	3.88	4.41	8	0.69	0.01	201 (62.0%)	0.04
				Friendly	324	3.71	3.88	4.41	8	0.69	0.01	201 (62.0%)	0.04
Richard III	0.21	1.98	0.70	All	74	4.45	2.37	2.88	6	0.69	0.06	50 (67.5%)	-0.10
				Hostile	31	1.67	2.08	3.25	4	0	0.05	10 (32.2%)	-0.20
				Friendly	68	4.12	2.45	2.98	6	0.69	0.06	45 (66.1%)	-0.03
Romeo & Juliet	0.34	1.8	0.80	All	404	6.10	2.76	3.32	7	0.82	0.02	398 (98.5%)	-0.33
				Hostile	144	2.33	2.93	5.88	7	0.17	0.02	131 (90.9%)	-0.36
				Friendly	385	5.67	2.84	3.43	7	0.84	0.01	350 (90.9%)	-0.32
Tempest	0.72	1.38	0.93	All	716	7.40	3.54	3.28	11	0.57	0.01	707 (98.7%)	-0.08
				Hostile	321	2.25	4.10	7.12	9	0	0.01	288 (89.4%)	-0.39
				Friendly	664	6.98	3.83	3.34	12	0.62	0.01	547 (82.3%)	0.10
Titus Andronicus	0.55	1.45	0.84	All	67	3.49	2.83	3.36	7	0.68	0.05	43 (64.2%)	0.01
				Hostile	324	3.71	3.88	4.41	8	0.69	0.01	201 (62.0%)	0.04
				Friendly	324	3.71	3.88	4.41	8	0.69	0.01	201 (62.0%)	0.04
Troilus & Cressida	0.29	1.69	0.87	All	74	4.45	2.37	2.88	6	0.69	0.06	50 (67.5%)	-0.10
				Hostile	31	1.67	2.08	3.25	4	0	0.05	10 (32.2%)	-0.20
				Friendly	68	4.12	2.45	2.98	6	0.69	0.06	45 (66.1%)	-0.03
Twelfth Night	0.69	1.23	0.79	All	404	6.10	2.76	3.32	7	0.82	0.02	398 (98.5%)	-0.33
				Hostile	144	2.33	2.93	5.88	7	0.17	0.02	131 (90.9%)	-0.36
				Friendly	385	5.67	2.84	3.43	7	0.84	0.01	350 (90.9%)	-0.32

a)

b)

Table 2. Topological measures used to assess the realism of the character networks of: a) Shakespeare’s plays [232]; b) mythological narratives [154].

The approach adopted in these articles is globally the same: they describe their character networks using standard topological descriptors (e.g. degree distribution, average distance, transitivity, cf. Section 6.1), as illustrated by Table 2, and compare their counterparts obtained for real-world and/or random networks. They generally find both similarities and differences. For instance, the *Marvel* network possesses certain properties observed in real-world networks: presence of a giant component, small-world (small average distance and high transitivity), scale-free (power law-distributed degree). However, it also presents important differences: its transitivity is much lower, and its degree is much smaller, reflecting the fact that *Marvel* characters tend to collaborate more often with the same people than real-world agents do [8]. This can be explained by legal matters related to the *Comics Authority Code* [92]. In the case of myths, *parts* of the networks are realistic whereas others are not. For instance, the Old English epic poem *Beowulf* is modeled by a relatively realistic network, provided one ignores the eponymous character: this is likely caused by the concentration of fantastic plot elements on him, whereas the other characters are based on real people [121, 122, 154, 157].

6.2.3 Role Detection. A number of authors leverage character networks for role detection in plays, novels and movies. i.e. assigning one among several predefined narrative roles to each character. A consensual definition would be that a character plays a role in the story in the sense that it follows “characterization archetypes” [206], such as being a protagonist or an antagonist. Once detected, roles can be used to solve higher-level problems such as plot classification (Section 6.2.4), storyline detection (Section 6.2.5), and plot summarization (Section 6.2.6). Identifying the protagonists can help characterize the structure of the plot. For instance, in his series *Les Rougon-Macquart*, Zola applies a technique called *polyfocalisation* and consisting in switching the narrative focus between several characters, which are therefore all protagonists [206, 207]. On the contrary, Rousseau’s *Les confessions* is largely focused on a specific character.

Based on this assumption, several authors take advantage of the network structure to predict some of these traits. Most of them use a supervised approach: Suen *et al.* train a classifier to distinguish scripts based on various criteria (e.g. cinema vs. theater, publication year) [238]; Hettinger *et al.* classify German novels in terms of genre [105]; Li *et al.* do the same for plays [145]; Holanda *et al.* try to distinguish texts according to three classes (pure fictions, pure biographies, and legends –a mix of both) [108]. A few authors adopt a non-supervised, more exploratory approach, through clustering: Ardanuy & Sporleder group similar novels and study the uniformity of such groups in terms of genre and writer [11, 12]; Waumans *et al.* want to gather episodes of the same series. Rochat & Triclot [209] manually classify works in four plot classes (heroic-core, unicentered, acentric, polycentric).

The general approach consists in extracting the character network of the considered fiction work, before computing a collection of topological measures used as features by the classifier: network size, density, transitivity, degree distribution, diameter, radius, centrality measures, and others. Li *et al.* compare these to their weighted generalization of graph motifs [145]. Most authors additionally use meta-data (medium type, genre, writer, date, length/duration) and content-related features (point of view of the narrator, word frequency, topics). Ardanuy & Sporleder [11, 12] also leverage the plot dynamics: they split each narrative into a few predefined phases (exposition, rising action, climax...) and separately compute the features for each of them.

In the end, the nature of the most discriminative features largely depends on the considered data and classification task. For instance, Suen *et al.* [238] observe that the best features to distinguish between plays and movies are degree-related, because plays generally have a single largely dominating character, whereas movies possess several of them. When discriminating between fictions and biographies, Holanda *et al.* [108] remark that the most discriminant feature is the number of *hapax legomena* (single-occurrence characters), which increases with the level of realism of the book. They assume that novel writers tend to use characters several times once they have made the effort of introducing them, whereas character occurrences are dictated by historical events in biographies.

6.2.5 Story Decomposition. Character networks are used in the literature to break down a story into its constitutive parts. Solving this problem is relevant for numerous information retrieval tasks, such as plot summarization or comparison, and can be expressed under various forms.

Storyline detection consists in identifying the multiple subplots intertwined by the writer/director to constitute the plot of his narrative. The methods proposed in the literature are applied to movies [191, 267] and TV series [269]. They postulate that each storyline is organized around a leading character. They first use role detection to detect these main characters, and then community detection to determine their entourage. Each scene is then associated to one of the communities depending on the involved characters, and storylines are obtained by ordering them chronologically. Scenes involving characters from multiple communities are points of contact between the storylines: they generally correspond to key moments in the plot, and are therefore particularly interesting for higher-level tasks such as automatic summarization. Certain authors observe that main characters are likely to appear in multiple subplots, and therefore assume that they can belong to several communities [191, 268].

Story segmentation constitutes another form of story decomposition. It consists in splitting the plot into consecutive and meaningful phases. When dealing with movies and TV series, this amounts to gather shots constituting scenes or substories (so-called *sequences*, cf. Section 4.1.1). Traditional methods focus only on low-level audiovisual features [268], but recent approaches take advantage of the plot-related semantics encoded in character networks. Weng *et al.* [269] define a character similarity measure based on neighborhood proximity in the character network, then

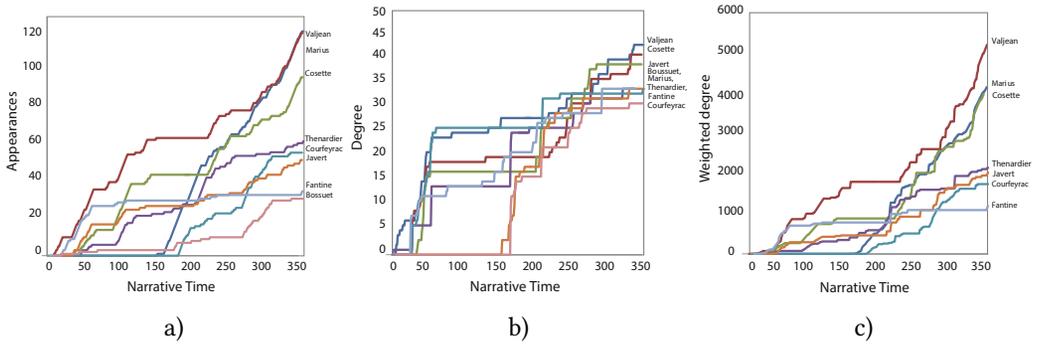


Fig. 19. Evolution of the cumulative dynamic network of *Les Misérables* used in [167] to detect plot phases: a) number of occurrences of the characters; b) degree; c) strength.

use it to compute scene similarity based on the similarity of the involved characters. The plot is then split by identifying strongly dissimilar consecutive scenes. Min & Park [167] want to detect standard plot phases (exposition, rising action, climax, falling action, resolution) in novels. For this purpose, they use a dynamic cumulative network and detect significant increases of vertices and edges, as well as stable periods, to identify these phases (see Figure 19). They subsequently use sentiment analysis and topic detection to improve their tool by taking the content into account.

6.2.6 Summarization. Several approaches use character networks to summarize audiovisual narratives, in place or in addition to the traditional low-level audiovisual features. This task consists in producing an extractive summary by combining important scenes picked from the original narrative.

Tran *et al.* [65, 244, 246] propose a method for the automatic summarization of movies. They first extract a co-occurrence network, and use a combination of standard topological measures to identify the main characters. They then assign a so-called *social score* to each scene depending on the nature of the involved characters. They combine this score with video-related features such as inter-scene distance and duration, to identify the scenes of particular interest, which are finally used to generate the summary. Bost *et al.* [33, 34] propose an extractive method to summarize TV series seasons, from the perspective of a character of interest (by opposition to summarizing the whole plot). They extract a dynamic conversational network using narrative smoothing (cf. Section 5.2.2), in order to take into account the non-linear nature of the narratives (eg. parallel subplots). Their assumption is that important events occurring in the storyline of characters are associated to durable changes in their social environment. They use the vertex neighborhood as a proxy to such social environment, and detect these events by clustering the sequence representing the successive states of this neighborhood. Their algorithm combines this graph-based approach to a more traditional analysis of the filmmaking grammar to generate the character-oriented summaries.

Tsai *et al.* [252] use a network of character groups, instead of individual characters (cf. Section 5.1.1): each vertex represents possibly overlapping subsets of characters, and edges between them model inclusion relations involving these subsets. The authors assume that a movie is typically built by first exposing the characters in scenes involving few of them, before developing the proper story by making more of them interact during more crowded scenes. They use an incoming degree-based centrality measure to identify the most important character groups in the plot and partition the network around them. The video summary is finally built by solving an optimization

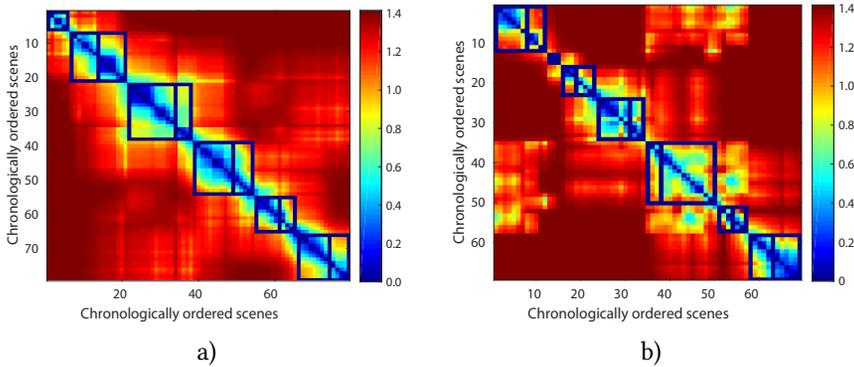


Fig. 20. Evolution of the vertex neighborhood, and segmentation of the narrative, for two characters from the *Game of Thrones* TV series [33]: a) JAIME LANNISTER ; and b) ARYA STARK.

problem that consists in selecting the scenes involving the most central vertices (character groups), under some user-specified length constraints.

6.2.7 Other Applications. In the literature, a number of less studied problems are also solved through the analysis of character networks: we mention them briefly in this section.

Improvement of pre-processing. Some authors leverage character networks to improve the execution of tasks which are considered as pre-processing in our extraction process. As a part of their method to perform scene segmentation in movies, Xu *et al.* [148] separately extract two co-occurrence networks: one based on the movie video, one based on its script. These networks are used to match faces from the video to character names in the script.

Yeh & Wu [274] propose an iterative approach taking advantage of a co-occurrence networks to perform face clustering in movies. They start by applying a generic face detection method to obtain a raw estimation of the face clusters. They use their own method [273] to extract a co-occurrence network, assuming each face cluster corresponds to a character. Similar clusters are merged based on features extracted from both the network and the video. The resulting clusters are used to extract a new co-occurrence network, and the same process is iteratively repeated to improve the quality of both face clusters and character network.

Generation and prediction of narratives. Other authors use character networks to predict or generate stories. Sack [214–216] proposes a method to produce proto-narratives, i.e. sequences of events involving characters, useful to later build stories. Starting from a randomly generated unbalanced signed network, it iteratively switches one edge sign in a randomly picked unbalanced triangle, thus making it balanced (cf. Figure 21a). The whole network ultimately becomes balanced. Each sign switch is considered as an event in the proto-narrative: either befriending (negative edge becoming positive) or betraying (vice versa). The number of switches undergone by a vertex can be used *a posteriori* to determine the type of behavior of the corresponding character.

Nalisnick & Baird have a similar approach, when they try to use structural balance to predict the evolution of Shakespeare’s plays [176]. However, they use Marvel *et al.*’s model [162], which consists in repeatedly squaring the signed adjacency matrix.

Recommendation systems. Lee & Jung [144] propose a recommendation system, allowing to propose a movie to a user depending on the ones he already likes. The authors first extract a so-called *affective* dynamic co-occurrence network, which is basically a signed network representing

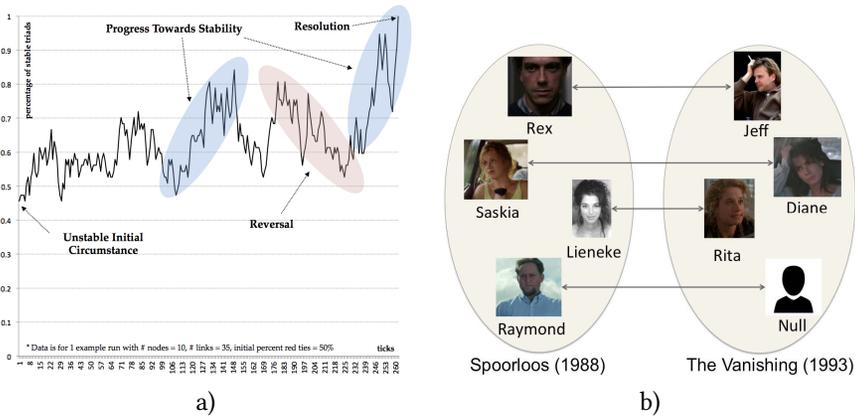


Fig. 21. Evolution of the proportion of stable triads, used in [215]; b) Alignment between the characters of two different movies (*Spoorloos* and *The Vanishing*) [50].

the friendly vs. hostile relationships between characters. Based on the evolution of the centrality of the vertices and the relationships between them, they identify significant affective changes, which are assumed to correspond to major plot events. They then build a lattice-based representation of the plot using these events. These are finally used as proxies to compare plots.

Alignment of narratives. In [50] Chaturvedi *et al.* aim at automatically detecting movie remakes. They do not work with the videos themselves, but rather with textual summaries obtained from Wikipedia. They define two similarity measures between the narratives: one based on the plot and the other on the characters. The former is the cosine similarity applied to bag-of-words representations of the events and entities detected in the summaries. The latter is the average character alignment between the narratives. This alignment is computed by matching the characters of two distinct narratives (cf. Figure 21b) depending on their name, gender, prominence in the narrative, and relationships with other characters (the only character network-based aspect of the process). The authors combine both measures to train a classifier into predicting so-called *remake clusters* (classes containing a movie and all its remakes).

Elsner [69] also works with texts, but these are 19th century romantic novels. In order to compare their narratives, he defines a character similarity measure relying on both individual and relational aspects. The former include the characters' prominence as well as lexical properties of the text surrounding the characters' occurrences, represented as unigram distributions. Relationships are detected based on a variant of the co-occurrence approach described in Section 4.1 of the *Main Article*, and Elsner uses the resulting edge weights when comparing characters. The characters are first aligned based on the individual features. Then, the relational similarity between two characters belonging to two distinct narratives depends on how much the edge weights of their aligned neighbors match. In addition, Elsner takes the narrative dynamics into account by considering how all these aspects evolve over the narrative. He then trains a classifier to distinguish between actual novels and randomly generated ones.

7 DISCUSSION, OPPORTUNITIES AND PERSPECTIVES

Our review of the literature reveals that there are still a number of issues to solve and directions to explore, at each of the steps constituting the network extraction (Section 7.1) and exploitation (Section 7.2) processes.

7.1 Network Extraction

We identify three main types of open problems regarding the extraction of character networks: the improvement and development of tools to identify characters and interactions in the various types of media used to create narratives (Section 7.1.1); of approaches to construct dynamic character networks based on this information (Section 7.1.2); and of evaluation methods to assess the performance of these identification and construction tools (Section 7.1.3).

7.1.1 Character and Interaction Identification. It appears from this survey that all media are not equal regarding the character and interaction identification processes. This is because they rely on the resolution of a number of medium-specific lower-level problems, for which the state-of-the-art performance is not always as advanced in all domains.

Textual Narratives. With textual narratives, the literature shows that automatic methods work relatively well when dealing with the simplest tasks: character occurrences can be identified through NER [95], and detecting their co-occurrences is quite straightforward. A number of authors extract character networks based on these tools only. However, there is still work to do in order to leverage more of the information conveyed by text. Character unification is not efficiently solved, as it requires detecting all forms of character aliases, and most of all performing anaphora resolution, which is a difficult NLP problem [194]. There are also research opportunities with the identification of advanced types of interactions between characters, besides simple co-occurrences. This task is much harder, as it requires analyzing the semantics of the text, for instance by detecting action verbs and their respective subjects and objects. For all these open problems, as explained in the survey, the specific characteristics of fiction offer both additional difficulties and leverage compared to non-fiction content.

Audiovisual Narratives. By comparison with text, the results obtained with automatic tools on audiovisual narratives are much less satisfying, to the point where many authors perform certain tasks such as face recognition manually. This difference might explain why there are many more articles dealing with text than audiovisual narratives. The low-level tasks necessary to detect character occurrences (face detection and tracking, speaker segmentation) and unify them (face track clustering, speaker clustering) are still open problems when performed in such uncontrolled conditions. Regarding interaction detection, in practice authors basically focus on co-occurrences only, as identifying more precise interactions requires solving even lower-level problems. For instance, to get a conversational network, one has to perform lip motion detection or speaker identification to determine who is speaking. Furthermore, interlocutors may be tricky to detect when multiple characters are involved in the same scene. Obtaining an action-based network would require determining which character performs which action on which other character [262]: there is no trace of such attempt in the literature.

It seems difficult to automate the whole character network extraction pipeline as long as all these basic processing steps cannot be efficiently performed. Moreover, unlike for textual narratives, here the specific properties of fiction result in significant difficulties compared to non-fiction, and yield little additional leverage. On the bright side, multimodal approaches constitute a very promising perspective, as some authors try to combine information extracted from the video and audio streams, but also from text (scripts, transcriptions, external sources such as wiki pages).

Comics. This might be surprising, but comics appear to be even harder to process as a narrative than video content. This is because the information they convey can take a number of different forms (text, drawings, balloons, panels, onomatopoeia, effect lines, textures), each one needing some specific processing. Moreover, the medium itself is not very normalized, in the sense that

distinct artists may express the same ideas in very different graphical ways. As a result, all the (few) authors extracting character networks from comics operate manually [8, 92, 209].

Detecting and unifying character occurrences in comics require solving certain problems similar to those met for videos, such as face detection, but in an even more adverse context (non-anthropomorphic and/or highly deformed characters). Moreover, they also require performing specific lower-level tasks such as panel segmentation or bubble detection, which are still open problems. For interaction detection, like for audiovisual narratives, co-occurrences are easy to get once the characters have been identified (provided the panels have been properly segmented), but identifying more precise interactions is hard. Leveraging conversations would essentially depend on the extraction of text and locating of bubbles and captions, two difficult tasks which are not satisfactorily tackled by state-of-the-art methods. Identifying specific actions that characters perform on each other would require associating semantics to specific body postures, which is also an open problem. To finish with comics, it is worth noticing that we could not find any article dealing with *photonovels*: on the one hand, their economical interest is questionable, as they felt out of fashion, but on the other hand they might be easier than comics, as they rely on photographs instead of drawings, while retaining certain conventions of comics.

General Remarks. To summarize, independently from the type of medium, there are many low-level open problems regarding the extraction of characters and interactions in fiction works. Recently, deep learning methods have proven to be efficient on classic NLP and multimedia tasks (e.g. [87]), and are therefore a promising perspective to solve most of these problems. Alternatively, a promising change of perspective would be to break with the pipeline approach, which is currently the norm for character network extraction, as illustrated by Figure 1. A pipeline is a chaining of separate tools, each one aiming at solving a specific subproblem.

All approaches first attempt to detect characters before trying to capture interactions, possibly propagating errors made at the earlier stages of the process. Instead, both tasks could probably benefit from one another and be performed jointly, as demonstrated by Yeh & Wu for face clustering [274]. Another possible option to avoid or reduce the pipeline limitations is to leverage deep learning methods and adopt a supervised *end-to-end* approach over parts, or even the whole extraction process. However, deep learning methods require to be trained on large corpora, and current fiction-related corpora are small, or do not even exist for certain tasks (especially for comics). Thus, before being able to leverage deep learning methods, it is likely that the community will have to constitute or extend such corpora, a tedious work often performed collectively. An alternative would be to take advantage of existing corpora to perform *transfer learning*, based on existing models built on non-fiction data.

7.1.2 Dynamic Networks. Many authors using character networks to solve higher-level problems identify the dynamics of the story as a very important aspect to be taken into account. This is especially true in long-term narratives such as TV or novel series, in which relationships and characters are likely to change over time. Yet, the overwhelming majority of methods proposed in the literature rely on *static* networks. This may be because extracting *dynamic* networks requires making tough methodological choices such as choosing the size of the temporal window. Moreover, time can be modeled in several ways in a network [7, 109]. Finally, there are much fewer off-the-shelf tools to analyze dynamic networks, and these are more difficult to use.

However, this issue deserves to be explored, as dynamic networks are likely to improve both our understanding of the structure of fictional character networks, and the performance of the methods taking advantage of this structure to solve problems such as those presented in Section 6.2. It is even likely that such networks are necessary to solve certain intrinsically dynamic problems such as narrative generation (see Section 6.2.7).

7.1.3 Evaluation and Methodological Choices. As described before, extracting a reliable character network requires solving efficiently a number of media-specific problems (ex. NER, face detection), most of them difficult.

Evaluation Data and Methods. This raises several questions, the first one being that of *performance assessment*. As mentioned before, all existing extraction methods rely on the pipeline approach. Consequently, each processing step constituting the pipeline must be assessed separately. Performing such a task requires an appropriate ground truth, designed for the considered problem. Such annotated corpora already exist for most problems, but they are generally based on non-fiction data, in which case they do not account for the specific characteristics of fiction described throughout this survey. To obtain a reliable performance evaluation, the test dataset must be built on fiction, which ties this point to the issue of machine learning training that we mentioned before: for a number of problems, such corpora are still to be constituted. But it is also important to estimate the relevance of the extracted character networks, i.e. the output of the pipeline. Up to now, only a few authors have tried to conduct this type of assessment on character networks [4, 140]. This task necessitates ground-truth networks to which one can compare the networks extracted using the pipeline, so here too there is a significant manual annotation work to do. Besides these data-related aspects, some methodological questions remain to be solved. There are many possibly ways to compare two graphs [270], and it is not obvious which ones are the most appropriate to the case of character networks. In addition, this aspect could be dependent on the high-level use one wants to make of the network (ex. story decomposition, summarization).

Methodological Choices. Solving the data and methods issues described above is important not only to evaluate the performance of the processing steps constituting the extraction pipeline, but also to drive methodological choices. Indeed, there are a number of open questions regarding which approaches are the most appropriate to handle these steps, or if some steps should be handled at all, and assessing their effect on the overall pipeline performance is necessary to discriminate between them. For instance, in textual narratives, it is not clear whether detecting all forms of character mentions (including nominals) instead of just looking for proper nouns, and/or solving anaphoras instead of ignoring them, would significantly improve the quality of the extracted networks. The same goes when detecting interactions: we do not know if conversation-based networks are more informative than co-occurrence-based ones, yet they are much harder to extract. To generalize, we do not know whether solving such harder low-level problems would result in better character networks: this remains to be tested. Solving efficiently these problems require significant extra efforts, so this is a very important question, yet it is still open. Very few [68, 115] of the reviewed articles try to tackle the problem, by assessing the quality of the extracted networks. Solving the evaluation-related issues would also allow assessing the relevance of using dynamic networks instead of static ones, and/or to identify which parameter values are optimal (e.g. size of the narrative unit when detecting co-occurrences, or of the sliding window when building dynamic networks). It would also allow comparing the different ways of computing edge weights, as recent articles on real-world conversational networks have shown their importance for classification tasks [187]. Finally, it is important to note that the relevance of the extracted network is likely to depend greatly on the high-level task for which one plans to leverage this network.

7.2 Network Analysis

We distinguish between the development of methods aiming at describing character networks (Section 7.2.1), and the exploitation of character networks to solve high-level problems (Section 7.2.2).

7.2.1 Descriptive Tools.

Exploration/Adaptation of Existing Tools. The tools used in the literature to describe character networks are standard, very widespread complex network topological measures (see Section 6.1). Yet, the field has produced many more tools which are not as known (e.g. motif-based measures, backbone detection, core-periphery structures, community structure-related vertex roles), but still likely to be relevant to study character networks (see [60] for a review). Therefore, there is a large number of opportunities to explore here. Some of these tools (e.g. motif-based measures) are designed for large networks, which means they would be relevant only for networks representing collections of narratives, like the comic networks mentioned in the survey [8, 92]. Another interesting possibility is to adapt existing measures to the specific case of character networks. For instance, certain authors identify one or two characters as extremely central in the studied narrative (protagonist vs. antagonist). Then when computing distance-based centrality measures such as the closeness, why not focusing only on the shortest path joining the character of interest to the protagonist, instead of considering all pairs of characters?

Additional Information. Another promising perspective is to follow the evolution currently taking place in the field of complex network analysis, and to integrate more information in the network, in order to apply tools able to leverage this additional information. A few authors mentioned in the survey have started extracting such networks, by including time (dynamic networks), relationship polarity (signed networks), or individual information (attributed networks) in their model. However, they focus more on the extraction process than on the way to efficiently use this information. For instance, several approaches exist to detect significant changes in dynamic networks [7], which could be applied to identify plot twists, or to segment the narrative. Regarding signed networks, there are alternatives to the concept of structural balance, for instance relaxations allowing to identify mediator groups in a situation of conflict: this could be of interest when analyzing so-called *adversarial* movies [63]. Vertex attributes (see Section 3.4.2) can be the object of the description through measures such as assortativity, which allows quantifying how much homophilic the network is. But they can also help interpreting the output of other tools, for instance in the context of community detection, by associating certain attribute values to specific communities. To conclude, note that it is possible to combine all these aspects (time, attributes, polarity) in one network, which would require very specific descriptive tools.

Network Simplification. Another, more minor, open question, is that of network simplification. Most narratives contain a number of extras, i.e. very minor characters used only to strengthen the level of realism of the story, or set the atmosphere. They might be considered as noise though, as they are basically part of the set or background, and do not affect the story in any way. Certain authors try to remove such characters, for example by trimming leaves from the character network. This allows, in particular, easing the visual exploration of the network. However, the objective effect on the network topological properties and high-level problem at hand remains to be identified and measured, and some other methods might be more appropriate.

7.2.2 Applications.

Methods Transposition. It is worth noting that the articles using character networks to solve high-level problems are quite segregated depending on the medium of the considered narratives (textual vs. audiovisual). This makes sense for the low-level part of the network extraction process, as it is medium-specific, but not for the exploitation of the extracted networks, as these do not vary much depending on the narrative medium. Consequently, it seems possible to transpose the methods designed to solve certain problems on a given medium to another one. For instance, text-based networks have been used to assess literary theories, but we could not find any article trying to do the same for cinematographic ones, e.g. the widespread use of the so-called *Hollywood*

Formula to design movie plots. On the contrary, character networks have been used to generate summaries of TV series: the same general principle could be applied to novels.

Narrative Classification. A number of high-level applications taking advantage of character networks consist in solving some classification problem, such as grouping narratives depending on their genre, time of publication, or creator. A related task is that of narrative recommendation (e.g. movie recommendation), as it can also be formulated as a prediction problem. In both cases, all authors feed their classifier with a few topological measures, generally the same very standard measures mentioned in Section 7.2.1. But there is no reason why these specific measures would be more appropriate for a classification task than the other measures defined in the literature. Also, the most relevant measures are not necessarily the same depending on the considered classification task. A first improvement would therefore be to adopt a more exhaustive approach, and consider a much larger number of topological measures, picking them so that they characterize the network at various scales and scopes (see for instance [187] on real-world conversational networks). Another improvement would be to use representation learning instead of manually selecting discriminant features. This could be performed by using *graph embeddings* [58], a recent transposition to graphs of the NLP concept of *word embeddings* [43]. This approach allows directly learning the most appropriate numeric representation of the character network for the considered classification task. However, it requires access to a large amount of data.

Trans- and Cross-media. Recently, cross- and trans-media narratives have become popular, and their automatic processing constitutes another promising perspective. Crossmedia means that the same narrative is repeated over several different media, e.g. a novel and its movie adaptation. Transmedia means that different narratives belonging to the same fictional universe are expressed using different media, e.g. the sequel of a movie taking the form of a novel. This raises the interesting problem of *network alignment*, i.e. determining which character in one network corresponds to which character in the other (see also Section 6.2.7). In the case of transmedia, one could use character networks to look for plot intersections between the different narratives, and possibly detect divergences. They could also be used to help the user navigate among the sometimes numerous and complex narratives (e.g. super hero universes). In the case of crossmedia, adapting a narrative for a different medium generally involves changing the plot or even the story, and adding or removing characters. Identifying which changes a character network undergoes during such adaptation could help understanding and maybe automating this process. Moreover, having access to the same story under different forms could be leveraged to improve network extraction itself, by adopting and extending the multimodal approach already mentioned in Section 7.1.1.

Narrative Generation. Finally, automatic plot generation constitutes another promising perspective. Performing this task is likely to require leveraging some additional information, as mentioned in Section 7.2.1, in particular: dynamic networks, in order to represent existing plots and study them to model their typical evolution; and signed networks to deal with antagonistic stories, which are very frequent. A few recent articles deal with plot generation (see Section 6.2.7), but there is still much work to do. Designing models to generate networks which would be realistic according to some criteria of interest constitutes an important part of the complex networks field⁵, so this task could benefit from the work already conducted on this issue.

However, it is worth noting that if generating a sequence of events corresponding to a plot looks achievable in the short-term, automatically converting it into a proper narrative seems like a very difficult problem. It is possible to adopt an extractive approach, i.e. to build the narrative by selecting

⁵E.g. the famous Watts-Strogatz [263] and Barabási-Albert [17] models, which respectively produce small-world and scale-free networks

narrative bits among a predefined collection. The alternative is to generate the narrative outright, but this seems realistic only for certain media, in particular text, for which efficient language-based models exist.

A METHODS FOR THE EXTRACTION OF FICTIONAL CHARACTER NETWORKS

Table 3 shows the list of articles describing methods to extract character networks from works of fiction. Although many other articles are mentioned through the survey when dealing with specific points of the process described in Figure 1, in this table we only focus on methods aiming at extracting proper networks. Figures 22 and 23 display how these articles relate.

Resource Availability. We made these resources publicly available online⁶: the table as a spreadsheet file, the figures as separate files, the networks as Gephi and Graphml files, and the bibliographic entries as a BibTeX file.

Citation Network. Figure 22 shows the citation network of the articles listed in Table 3. Note that these articles also cite many bibliographic references not included in this table: we do not represent them in the figure, in order to improve its readability.

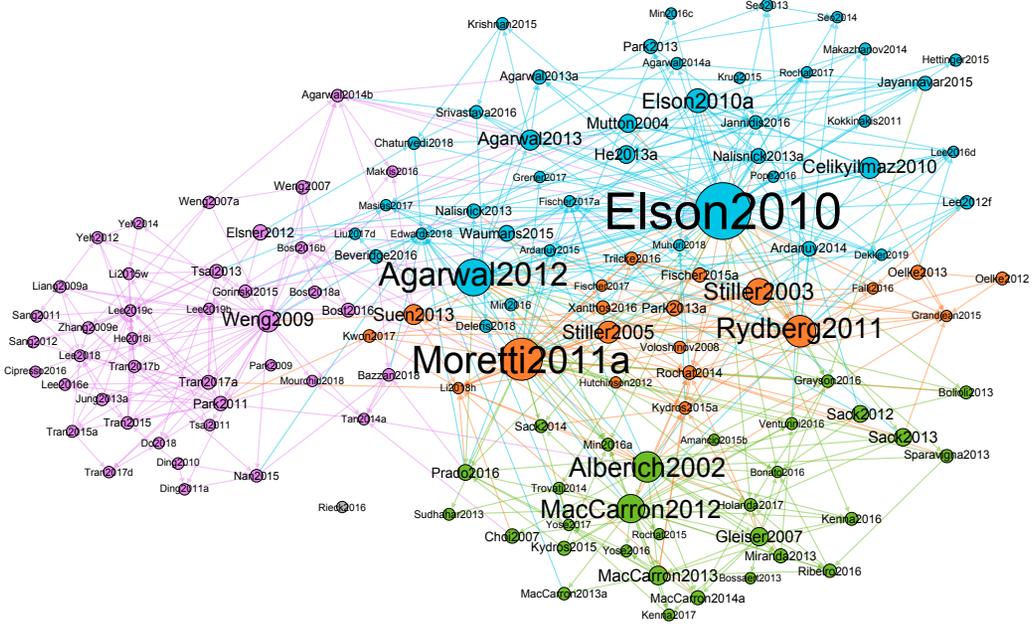


Fig. 22. Citation network for the articles of Table 3. Figure and data available at [10.6084/m9.figshare.7993040](https://doi.org/10.6084/m9.figshare.7993040) under CC-BY license.

Each vertex corresponds to an article and an edge represents the fact that an article cites another article. As is usually done for this type of network, edges are directed from the cited article to the citing one, in order to show the flow of ideas. Each vertex is identified by its BibTeX key, which corresponds to the key used in the BibTeX file publicly available online. The colors of the vertices represent the communities detected by the Louvain algorithm [24] (a widespread modularity optimization-based method). The size of the vertices are proportional to their Hub score [125],

⁶<https://doi.org/10.6084/m9.figshare.7993040>

Work of fiction	Ref.	Relationships					Graph				Application
		Cc.	Cv.	M.	Ac.	Af.	W.	Di.	S.	Dy.	
5 classic novels	[126]	Y	N	N	N	N	N	N	N	N	Benchmark for graph processing tools
Marvel universe	[8]	Y	N	N	N	N	N	N	N	N	Level of realism
10 Shakespeare's plays	[231, 232]	Y	N	N	N	N	N	N	N	N	Level of realism
All of Shakespeare's plays	[173]	N	Y	N	N	N	Y	N	N	N/Y	Visualization
Dictionary of Greek and Roman mythology	[53]	N	N	Y	N	N	N	N/Y	N	N	Descriptive analysis
Marvel universe	[92]	Y	N	N	N	N	Y	N	N	N	Level of realism
Hollywood movies, and TV series	[267–269]	Y	N	N	N	N	Y	N	N	N	Storyline identification/segmentation
9 classical plays	[261]	N	Y	N	N	N	N	N	N	N	Level of realism
Movies and TV series	[190, 191]	N	Y	N	N	N	Y	Y	N	N	Role detection, story segmentation
Friends TV show	[279, 280]	Y	N	N	N	N	Y	N	N	N	Method assessment
15 Hollywood movies	[148, 217, 218, 282]	Y	N	N	N	N	Y	N	N	N	Face-name matching, Scene segmentation
Austen's <i>Pride & prejudice</i> , and <i>Emma</i>	[45]	N	Y	N	N	N	Y	N	N	N	Method assessment
10 adversarial movies	[63]	Y	N	N	N	N	Y	N	Y	N	Method assessment
300 19th century Swedish novels	[127]	Y	N	N	Y	Y	N	N	N	N	Method assessment
60 19th century British novels	[70–72]	N/Y	N/Y	N/Y	N	N	Y	N	N	N	Check literary theories
20 Hollywood movies	[64]	Y	N	N	N	N	Y	N	Y	Y	Community detection
Shakespeare's <i>Hamlet</i> , <i>Macbeth</i> , and <i>King Lear</i>	[170]	N	Y	N	N	N	N	N	N	N	Discussion about the characters
Collection of Greek Tragedies	[213]	N	Y	N	N	N	N	Y	N	N	Plot structure comparison
8 Hollywood movies	[253]	Y	N	N	N	N	N	N	N	Y	Scene segmentation
Carroll's <i>Alice in Wonderland</i>	[1, 3–5]	N	N	N	Y	N	N	N/Y	N	N/Y	Role detection
41 19th century novels	[69]	Y	N	N	N	N	Y	N	N	Y	Genre comparison
Rowling's <i>Harry Potter</i> novels	[110]	Y	N	N	N	N	Y	N	N	N	Method assessment
The <i>Old Testament</i>	[140]	N	Y	N	Y	Y	N	N	N	N	Method assessment
4 European tales	[121, 122, 154–157]	Y	Y	Y	Y	Y	N/Y	N	Y	N	Level of historicity
19th and 20th century Swedish novels	[184, 185]	Y	N	N	N	N	Y	N	N	N/Y	Visualization
6 canonical European novels	[214–216]	Y	N	N	N	N	Y	N	N	N	Proto-narrative generation
Frankel's <i>The Devil wears Prada</i>	[273, 274]	Y	N	N	N	N	Y	N	N	N	Face clustering
Manzoni's <i>I promessi sposi</i>	[26]	N/Y	N/Y	N	N	N	Y	N	N	N	Method assessment
Rowling's <i>Harry Potter</i> novels	[32]	N	N	N	Y	N	N	Y	N	N	Study of peer support
3 classic European novels	[102, 159]	N	Y	N	N	Y	N	N	N	N	Method assessment
10 Hollywood movie scripts	[118]	N	Y	N	N	N	Y	Y	N	N	Role detection
Homer's <i>Odyssey</i>	[168]	Y	Y	Y	Y	Y	N	N	N	N	Level of historicity
Shakespeare's plays	[175, 176]	N	Y	N	N	N	Y	Y	Y	Y	Method assessment
20 novels	[188, 189, 221, 222]	Y	N	N	N	N	Y	N	N	N	Descriptive analysis
Rowling's <i>Harry Potter and the Philosopher's Stone</i>	[228]	N	Y	Y	Y	Y	N	N	N	N	Level of realism
3 story books	[236]	N	N	N	Y	N	Y	Y	N	N	Role detection
173 theater and 580 movie scripts	[238]	N/Y	N/Y	N	N	N	Y	N	N	N	Classification of works
12 Hollywood movies	[252]	Y	N	N	N	N	N	N	N	N	Video summarization
674 movie scripts	[2]	N	Y	N	N	N	N	N	N	N	Performance evaluation
238 novels	[11, 12]	Y	N	N	N	N	Y	N	N	N	Classification of works
Tales and myths	[158]	Y	Y	Y	Y	Y	N	N	Y	N	Level of historicity, Categorization
Rousseau's <i>Les confessions</i>	[206, 208]	Y	N	N	N	N	N/Y	N	N	N/Y	Character roles
Star Trek and Star Gate movies and TV series scripts	[242]	Y	N	N	N	N	Y	N	N	N	Comparative study
Knowles' <i>The Legends of King Arthur and his Knights</i>	[251]	N	N	N	Y	N	N	N	Y	N	Method assessment
20 novels	[10]	Y	N	N	N	N	Y	N	N	N	Descriptive analysis
1,276 movie scripts	[94]	Y	N	N	N	N	Y	N	N	N	Script summarization
1,682 16-20th century German novels	[105]	Y	N	N	N	N	N	N	N	N	Classification of works

Work of fiction	Ref.	Relationships					Graph				Application
		Cc.	Cv.	M.	Ac.	Af.	W.	Di.	S.	Dy.	
60 19th century British novels	[115]	N	Y	N	Y	N	N/Y	N/Y	N	N	Check literary theories
617 movie scripts	[129]	N	Y	N	N	N	N	N	Y	N	Infer formality level of interactions
58 German novels	[113, 130]	N/Y	N/Y	N	N	N	Y	N	N	N	Role detection
Homer's <i>Illiad</i>	[137]	Y	N	N	N	N	N	N	N	N	Descriptive analysis
Emirikos' <i>The Great Eastern</i>	[136]	Y	Y	Y	Y	Y	Y	Y	N	N	Descriptive analysis
6 Hollywood movies	[146]	Y	N	N	N	N	N	N	N	Y	Video summarization
183 <i>Friends</i> episodes	[177]	Y	N	N	N	N	Y	N	N	Y	Method assessment
<i>The Wire</i> script	[195]	N	Y	N	N	N	Y	N	N	N	Role detection
Zola's <i>Les Rougon-Macquart</i>	[207]	Y	N	N	N	N	Y	N	N	N	Plot structure comparison
First 6 <i>Star Wars</i> movies	[248]	Y	N	N	N	N	Y	N	N	N	Role detection
Corpus of Hollywood movies	[65, 244, 246]	Y	N	N	N	N	Y	N	N	N	Movie summarization
Large corpora of European plays	[77–79, 250]	Y	N	N	N	N	N	N	N	N	Check literary theories
46 fantasy novels, and Hugo's <i>Les misérables</i>	[264]	N	Y	N	N	N	N	N	N	Y	Classification of works
Martin's <i>A Storm of Swords</i>	[23]	Y	N	N	N	N	Y	N	N	N	Descriptive analysis
3 modern novels	[29]	Y	N	N	N	N	Y	N	N	N	Model fitting
3 TV series	[34–36]	N/Y	N/Y	N	N	N	Y	N	N	N/Y	Video summarization
Tarentino's <i>Pulp Fiction</i>	[55]	N	N	N	Y	N	Y	Y	N	N	Descriptive analysis
Edgeworth's <i>The Absentee</i>	[75]	N	Y	N	N	N	N	N	N	N	Check literary theories
9 19th century British novels	[98]	Y	N	N	N	N	Y	N	N	N/Y	Descriptive analysis
Chinese Buddhist Canon	[139]	N	Y	N	N	N	Y	N	N	N	Descriptive analysis
10 Hollywood movies	[141, 144]	Y	N	N	N	N	Y	N	N	Y	Recommendation system
501 movie scripts	[160]	Y	N	N	N	N	Y	N	N	N	Method assessment
Hugo's <i>Les misérables</i>	[165–167]	Y	N	N	N	N	N	N	N	N/Y	Story segmentation
<i>La chanson de Roland</i> , and <i>Alice in Wonderland</i>	[197]	Y	Y	Y	Y	Y	N	N	N	N/Y	Descriptive analysis
Tolkien's <i>Middle-Earth</i> novels	[199]	Y	Y	Y	Y	N	N	N	N	N	Descriptive analysis
All of Shakespeare's plays	[200]	Y	N	N	N	N	Y	N	N	N	Comparative study
CMU Movie Summary corpus	[229]	Y	N	N	Y	Y	N	N	Y	N	Method assessment
Homer's <i>Illiad</i>	[259]	N	N	N	Y	N	N	N	N	N	Visualization
Molière's <i>L'école des femmes</i>	[272]	Y	N	N	N	N	Y	N	N	Y	Visualization
<i>Ossian</i> corpus of Scottish epic poems	[278]	Y	Y	Y	Y	Y	N	N	Y	N	Level of historicity
4 gospels of the <i>New Testament</i>	[97]	Y	N	N	N	N	N	N	N	N	Comparative study
14 British Victorian novels	[99]	Y	N	N	N	N	N	N	N	N/Y	Visualization
9 fictionary, legendary and biographical books	[108]	N	N	N	Y	N	N	N	N	N	Level of historicity
Shakespeare's <i>Hamlet</i>	[135]	N	N/Y	N/Y	N	N	Y	Y	N/Y	N/Y	Descriptive analysis
<i>Game of Thrones</i> TV series (Seasons 1–6)	[151]	N	N	N	Y	N	N	N	Y	Y	Descriptive analysis
Shakespeare's <i>Romeo & Juliet</i>	[163]	N	Y	N	N	N	N/Y	N	N	N	Role detection
Corpus of science-fiction works	[209]	Y	N	N	N	N	Y	N	N	N	Plot structure comparison
17 Hollywood movies	[247, 249]	Y	N	N	N	N	Y	N	N	Y	Role detection
<i>Friends</i> TV series	[21]	N	Y	N	N	Y	N	N	N	Y	Check narrative theory
Summaries of 577 original movies and remakes	[50]	N	N	N	N	Y	Y	N	N	N	Plot structure comparison
First 2 seasons of <i>Friends</i>	[62]	N	Y	Y	N	N	Y	Y	N	N	Method assessment
<i>Friends</i> TV series	[68]	N/Y	N/Y	N/Y	N	N/Y	Y	N	N	N	Comparison of extraction methods
8 Hollywood movies	[103]	Y	N	N	N	N	Y	N	N	N	Role detection
65 classic plays	[145]	N	Y	N	N	N	Y	Y	N	N	Classification of works
2 Hollywood movie scripts	[171]	N	Y	N	N	N	N	N	N	N	Descriptive analysis
Tagore's <i>Raktakarabi</i> and <i>Muktadhara</i>	[172]	N	Y	N	N	N	Y	Y	N	N	Role detection
1 TV show and 3 movies	[153]	Y	Y	N	N	N	Y	N	N	N	Method assessment

Work of fiction	Ref.	Relationships					Graph				Application
		Cc.	Cv.	M.	Ac.	Af.	W.	Di.	S.	Dy.	
Medieval Irish text <i>Cogadh Gaedhel re Gallaibh</i>	[277]	Y	Y	Y	Y	Y	N	N	N/Y	N	Level of historicity
20 classic and 20 modern novels	[61]	Y	N	N	N	N	Y	N	N	N	Descriptive analysis
12 Hollywood movies	[143]	Y	N	N	N	N	Y	N	N	Y	Method assessment
No data (methodological paper)	[142]	Y	N	N	N	N	Y	N	N	Y	Storyline identification
<i>Game of Thrones</i> TV series	[230]	N	N	N	N	Y	N	N	N	N	Link prediction

Table 3. List of methods designed to extract character networks from fictional works. The *Relationships* column indicates whether the edges correspond to co-occurrences (Cc.), conversations (Cv.), mentions (M.), direct actions (Ac.), or affiliations (Af.). The *Graph* column shows whether the extracted networks are weighted (W.), directed (Di.), signed (S.), and dynamic (Dy.). When the authors experiment with different methods, several values may appear in the same cell. Note that this table has been completed after the publication of the official ACM CS article. Table and data available at [10.6084/m9.figshare.7993040](https://doi.org/10.6084/m9.figshare.7993040) under CC-BY license.

a generalization of the degree designed to take higher order influence into account, in directed networks. It focuses on the outgoing edges, i.e. on how influential an article is.

The community structure shows the existence of 4 groups of articles with dense internal edges but comparatively sparsely connected to the other groups. They can be interpreted as groups of authors working in a relatively independent way, and possibly unaware of the work conducted in other communities. The purple community located in the left-hand part contains almost only articles dealing with video narratives, and almost all of them. Only a few articles from this community cite articles handling textual narratives, which highlights the lack of communication between these fields. The remaining communities are built around a core of several articles written by the same team: MacCarron *et al.* (green community), Elson *et al.* + Agarwal *et al.* (blue community), Rochat *et al.* + Moretti *et al.* (orange community). The nature of the extracted character network (co-occurrence, conversational, etc.) does not seem to be related to the community. Rather, they seem to correspond to the high-level problem tackled by the concerned articles. For instance, articles in the green group tend to study the historical relevance of certain texts, those in the orange one mainly originate from the field of literary analysis, and works from the blue group correspond to computer science methods aiming at improving the character extraction process. The article of Rieck & Leitte is completely isolated, as it does not cite any other article from the table, and is cited by none of them.

The centrality allows to identify a few important articles. The 2002 article of Alberich *et al.* [8] is a forerunner of the use of complex network analysis tools to study character networks, here for comics. In [170], Moretti makes the connection between graph theory and literary studies. Elson *et al.* are among the first to automate the extraction of conversational networks from novels in their 2010 article [71], and so are Agarwal *et al.* for action-based networks [3]. The 2012 article of MacCarron & Kenna [154] is the first of a long series of papers related to mythological works. In the video community, the 2009 paper of Weng *et al.* [269] describes the *RoleNet* method, which inspired many other approaches in the community.

Co-citation Network. Figure 23 shows the co-citation network of the articles from Table 3. Again, articles not listed in Table 3 are not represented. Each vertex corresponds to an article and the weighted edges represent how similar two articles are in terms of the articles they cite. To get this similarity value, we compute Jaccard's coefficient based on both sets of bibliographic entries contained in the two considered articles. Like in Figure 22, each vertex is identified by its BibTeX key, and its color represents its community as detected by Louvain. Unlike the previous network, the vertex size is proportional to its eigencentrality [28], as this network is not directed.

The community structure is constituted of 4 groups of articles relatively similar to those from Figure 22, as highlighted by the matching colors. Each community can be considered as a group of articles tending to cite the same set of papers, a set different from those of the other communities. The purple community corresponds to a group of articles dealing with video narratives, a bit like in Figure 22, except some of these articles are actually placed in other communities (e.g. Bost *et al.* our own previous work, in the orange group). Still, all 3 remaining communities mainly focus on textual narratives. The green one contains all the articles of MacCarron *et al.*'s team, which rely on manually extracted networks. But it seems its main characteristic is that these articles, as well as others in this community such as Miranda *et al.*'s 2013 article [168], study the level of historicity of narratives. Although the article of Rieck & Leitte is not directly connected to any other work from Table 3 in Figure 23, here it is placed in the green group, which means the article it mentions are mostly similar to the bibliography usually cited by those papers.

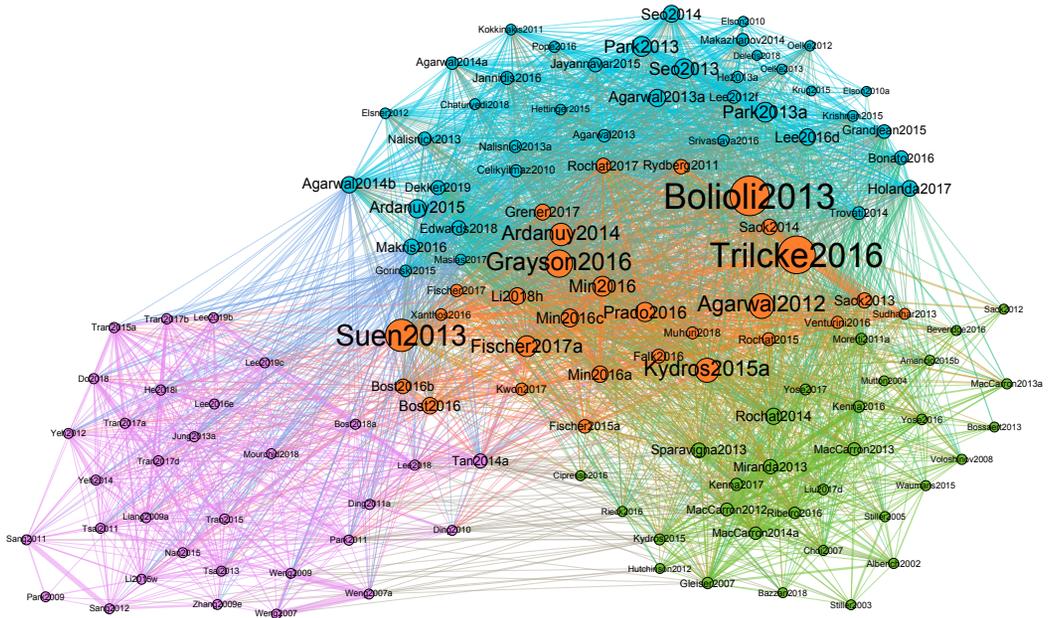


Fig. 23. Co-citation network for the articles of Table 3. Figure and data available at [10.6084/m9.figshare.7993040](https://doi.org/10.6084/m9.figshare.7993040) under CC-BY license.

REFERENCES

- [1] A. Agarwal, S. Balasubramanian, A. Kotalwar, J. Zheng, and O. Rambow. 2014. Frame semantic tree kernels for social network extraction from text. In *14th Conference of the European Chapter of the Association for Computational Linguistics*. <https://pdfs.semanticscholar.org/2b67/dfc9f9688eb2b53cdd9cd2ab3dcd373c82f0.pdf>
- [2] A. Agarwal, S. Balasubramanian, J. Zheng, and S. Dash. 2014. Parsing screenplays for extracting social networks from movies. In *3rd Workshop on Computational Linguistics for Literature*. ACL, 50–58. <http://www.aclweb.org/anthology/W14-0907>
- [3] A. Agarwal, A. Corvalan, J. Jensen, and O. Rambow. 2012. Social Network Analysis of Alice In Wonderland. In *NAACL - Workshop on Computational Linguistics for Literature*. http://www.cs.columbia.edu/~apoorv/Homepage/Publications_files/naacl2012.pdf
- [4] A. Agarwal, A. Kotalwar, and O. Rambow. 2013. Automatic Extraction of Social Networks from Literary Text: A Case Study on Alice in Wonderland. In *International Joint Conference on Natural Language Processing*. 1202–1208. http://www.aclweb.org/website/old_anthology/I13/I13-1171.pdf
- [5] A. Agarwal, A. Kotalwar, J. Zheng, and O. Rambow. 2013. Sinnet: Social interaction network extractor from text. In *International Joint Conference on Natural Language Processing - System Demonstrations*. 33–36. <http://www.aclweb.org/anthology/I13-2009>
- [6] A. Agarwal and O. Rambow. 2010. Automatic detection and classification of social events. In *Conference on Empirical Methods in Natural Language Processing*. 1024–1034. <https://dl.acm.org/citation.cfm?id=1870758>
- [7] C. Aggarwal and K. Subbian. 2014. Evolutionary Network Analysis: A Survey. *Comput. Surveys* 47, 1 (2014), 10. <https://doi.org/10.1145/2601412>
- [8] R. Alberich, J. Miro-Julia, and F. Rossetto. 2002. Marvel Universe looks almost like a real social network. *arXiv cond-mat.dis-nn (2002)*, cond-mat/0202174. <http://arxiv.org/abs/cond-mat/0202174>
- [9] A. Albright. 2015. The One With All The Quantifiable Friendships. <https://thelittledataset.com/2015/01/20/the-one-with-all-the-quantifiable-friendships/>. (01 2015).
- [10] D. R. Amancio. 2015. Network analysis of named entity interactions in written texts. *Europhysics Letters* 5, 114 (2015), 58005. <https://doi.org/10.1209/0295-5075/114/58005>
- [11] M. C. Ardanuy and C. Sporleder. 2014. Structure-based Clustering of Novels. In *3rd Workshop on Computational Linguistics for Literature*. 31–39. <http://www.aclweb.org/anthology/W14-0905>

- [12] M. C. Ardanuy and C. Sporleder. 2015. Clustering of novels represented as social networks. *Linguistic Issues in Language Technology* 12 (2015). <http://csli-lilt.stanford.edu/ojs/index.php/LiLT/article/view/60>
- [13] O. Augereau, M. Iwata, and K. Kise. 2018. A survey of comics research in computer science. *Journal of Imaging* 4, 7 (2018), 87. <https://doi.org/10.3390/jimaging4070087>
- [14] C. F. Baker, C. J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet Project. In *17th international conference on Computational linguistics*. 86–90. <https://doi.org/10.3115/980451.980860>
- [15] D. Bamman, B. O'Connor, and N. A. Smith. 2014. Learning latent personas of film characters. In *Annual Meeting of the Association for Computational Linguistics*. 352–361. <http://www.anthology.aclweb.org/P/P13/P13-1035.pdf>
- [16] N. Bansal, A. Blum, and S. Chawla. 2002. Correlation Clustering. In *43rd Annual IEEE Symposium on Foundations of Computer Science*. 238–247. <https://doi.org/10.1109/SFCS.2002.1181947>
- [17] A.-L. Barabási and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286, 5439 (1999), 509–512. <https://doi.org/10.1126/science.286.5439.509>
- [18] N. Bessiou, V. Moschou, and C. Kotropoulos. 2010. Speaker Diarization Exploiting the Eigengap Criterion and Cluster Ensembles. *IEEE Transactions on Audio, Speech, and Language Processing* 18, 8 (2010), 2134–2144. <https://doi.org/10.1109/TASL.2010.2042121>
- [19] J. Batchelor. 2018. Global games market value rising to \$134.9bn in 2018. <https://www.gamesindustry.biz/articles/2018-12-18-global-games-market-value-rose-to-usd134-9bn-in-2018>. (12 2018).
- [20] A. Bavelas. 1950. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America* 22, 6 (1950), 725–730. <https://doi.org/10.1121/1.1906679>
- [21] A. L. C. Bazzan. 2018. I will be there for you: six friends in a clique. *arXiv cs.SI* (2018), 1804.04408. <https://arxiv.org/abs/1804.04408>
- [22] K. Benzi. 2016. Exploring the Star Wars Universe (part 1). <http://kirellbenzi.com/blog/exploring-the-star-wars-expanded-universe/>. (01 2016).
- [23] A. Beveridge and J. Shan. 2016. Network of Thrones. *Math Horizons* 23, 4 (2016), 18–22. <https://doi.org/10.4169/mathhorizons.23.4.18>
- [24] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics* 2008, 10 (2008), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- [25] P. Boldi and S. Vigna. 2014. Axioms for centrality. *Internet Mathematics* 10, 3-4 (2014), 222–262. <https://doi.org/10.1080/15427951.2013.865686>
- [26] A. Bolioli, M. Casu, M. Lana, and R. Roda. 2013. Exploring the Betrothed Lovers. In *Workshop on Computational Models of Narrative*. 30–35. <https://doi.org/10.4230/OASiCs.CMN.2013.30>
- [27] P. Bonacich. 1972. Factoring and weighting approaches to status scores and clique identification. *Journal of Mathematical Sociology* 2, 1 (1972), 113–120. <https://doi.org/10.1080/0022250X.1972.9989806>
- [28] P. F. Bonacich. 1987. Power and centrality: A family of measures. *Amer. J. Sociology* 92, 5 (1987), 1170–1182. <https://doi.org/10.1086/228631>
- [29] A. Bonato, D. R. D'Angelo, E. R. Elenberg, D. F. Gleich, and Y. Hou. 2016. Mining and modeling character networks. In *International Workshop on Algorithms and Models for the Web-Graph (Lecture Notes in Computer Science)*, Vol. 10088. 100–114. https://doi.org/10.1007/978-3-319-49787-7_9
- [30] D. Bordwell and K. Thompson. 1993. *Film Art: An Introduction*. McGraw-Hill. <https://www.mhprofessional.com/9780073535104-usa-film-art-an-introduction>
- [31] C. Bornet and F. Kaplan. 2017. A Simple Set of Rules for Characters and Place Recognition in French Novels. *Frontiers in Digital Humanities* 4 (2017), 6. <https://doi.org/10.3389/fdigh.2017.00006>
- [32] G. Bossaert and N. Meidert. 2013. "We Are Only as Strong as We Are United, as Weak as We Are Divided" a Dynamic Analysis of the Peer Support Networks in the Harry Potter Books. *Open Journal of Applied Sciences* 3 (2013), 174–185. <https://doi.org/10.4236/ojapps.2013.32024>
- [33] X. Bost. 2016. *A storytelling machine ? : automatic video summarization : the case of TV series*. PhD Thesis. <https://tel.archives-ouvertes.fr/tel-01637270>
- [34] X. Bost, S. Gueye, V. Labatut, M. Larson, G. Linares, D. Malinas, and R. Roth. 2019. Remembering Winter Was Coming: Character-oriented Video Summaries of TV Series. *To appear in Multimedia Tools and Applications* (2019).
- [35] X. Bost, V. Labatut, S. Gueye, and G. Linares. 2016. Narrative smoothing: dynamic conversational network for the analysis of TV series plots. In *Advances in Social Networks Analysis and Mining - 2nd International Workshop on Dynamics in Networks*. 1111–1118. <https://doi.org/10.1109/ASONAM.2016.7752379>
- [36] X. Bost, V. Labatut, S. Gueye, and G. Linares. 2018. Extraction and analysis of dynamic conversational networks from TV series. In *Social Network Based Big Data Analysis and Applications*. Springer, Chapter 3, 55–84. https://doi.org/10.1007/978-3-319-78196-9_3
- [37] X. Bost and G. Linares. 2014. Constrained speaker diarization of TV series based on visual patterns. In *IEEE Spoken Language Technology Workshop*. 390–395. <https://doi.org/10.1109/SLT.2014.7078606>

- [38] X. Bost, G. Linarès, and S. Gueye. 2015. Audiovisual speaker diarization of TV series. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. 4799–4803. <https://doi.org/10.1109/ICASSP.2015.7178882>
- [39] M. Boudourides and S. Lenis. 2015. Network Analysis of Shakespeare’s MacBeth. <http://mboudour.github.io/2015/10/28/Shakespeare's-Macbeth-Network.html>. (10 2015).
- [40] M. Boudourides and S. Lenis. 2016. The Network and Trajectories of Transitions among Sentential Co-Occurrences of Characters of Arthur Conan Doyle’s A Study in Scarlet. <http://mboudour.github.io/2016/04/17/Arthur-Conan-Doyle's-A-Study-in-Scarlet-Network-&-Trajectories.html>. (04 2016).
- [41] H. Bredin and G. Gelly. 2016. Improving speaker diarization of TV series using talking-face detection and clustering. In *ACM international conference on Multimedia*. 157–161. <https://doi.org/10.1145/2964284.2967202>
- [42] A. D. Broido and A. Clauset. 2018. Scale-free networks are rare. *arXiv physics.soc-ph* (2018), 1801.03400. <https://arxiv.org/abs/1801.03400>
- [43] J. Camacho-Collados and M. T. Pilehvar. 2018. From Word To Sense Embeddings: A Survey on Vector Representations of Meaning. *Journal of Artificial Intelligence Research* 63 (2018), 743–788. <https://doi.org/10.1613/jair.1.11259>
- [44] D. Cartwright and F. Harary. 1956. Structural balance: A generalization of Heider’s theory. *Psychological Review* 63 (1956), 277–293. <https://doi.org/10.1037/h0046049>
- [45] A. Celikyilmaz, D. Hakkani-Tur, H. He, G. Kondrak, and D. Barbosa. 2010. The Actor-Topic Model for Extracting Social Networks in Literary Narrative. In *NIPS Workshop: Machine Learning for Social Computing*. <https://webdocs.cs.ualberta.ca/~denilson/files/publications/nips2010.pdf>
- [46] A. Chaganty and G. Muzny. 2015. *Quote Attribution for Literary Text with Neural Networks*. Technical Report. Stanford University. <http://cs224d.stanford.edu/reports/ChagantyArun.pdf>
- [47] S. Chaturvedi. 2016. *Structured Approaches for Exploring Interpersonal Relationships in Natural Language Text*. PhD Thesis. University of Maryland. <https://doi.org/10.13016/M28Z0S>
- [48] S. Chaturvedi, M. Iyyer, and H. Daumé. 2017. Unsupervised Learning of Evolving Relationships Between Literary Characters. In *31st AAAI Conference on Artificial Intelligence*. 3159–3165. <https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/viewPaper/14564>
- [49] S. Chaturvedi, S. Srivastava, H. Daumé, and C. Dyer. 2016. Modeling evolving relationships between characters in literary novels. In *30th AAAI Conference on Artificial Intelligence*. 2704–2710. <https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/viewPaper/12408>
- [50] S. Chaturvedi, S. Srivastava, and D. Roth. 2018. Where Have I Heard This Story Before? Identifying Narrative Similarity in Movie Remakes. In *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Vol. 2. 673–678. <https://aclanthology.coli.uni-saarland.de/papers/N18-2106/n18-2106>
- [51] H. Y. Chen and J. D. Choi. 2016. Character Identification on Multiparty Conversation: Identifying Mentions of Characters in TV Shows. In *17th Annual SIGdial Meeting on Discourse and Dialogue*. 90–100. <http://www.aclweb.org/anthology/W16-3612>
- [52] H. Y. Chen, E. Zhou, and J. D. Choi. 2017. Robust Coreference Resolution and Entity Linking on Dialogues: Character Identification on TV Show Transcripts. In *21st Conference on Computational Natural Language Learning*. 216–225. <http://www.aclweb.org/anthology/K/K17/K17-1023.pdf>
- [53] Y.-M. Choi and H.-J. Kim. 2007. A directed network of greek and roman mythology. *Physica A* 382, 2 (2007), 665–671. <https://doi.org/10.1016/j.physa.2007.04.035>
- [54] W.-T. Chu and W.-W. Li. 2017. Manga FaceNet: Face Detection in Manga based on Deep Neural Network. In *ACM on International Conference on Multimedia Retrieval*. 412–415. <https://doi.org/10.1145/3078971.3079031>
- [55] P. Ciproso and G. Riva. 2016. Computational Psychometrics Meets Hollywood: The Complexity in Emotional Storytelling. *Frontiers in Psychology* 7 (2016), 1753. <https://doi.org/10.3389/fpsyg.2016.01753>
- [56] A. Clauset and N. Eagle. 2007. Persistence and periodicity in a dynamic proximity network. In *DIMACS Workshop on Computational Methods for Dynamic Interaction Networks*. <http://arxiv.org/pdf/1211.7343>
- [57] P. Clément, T. Bazillon, and C. Fredouille. 2011. Speaker diarization of heterogeneous web video files: A preliminary study. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. 4432–4435. <https://doi.org/10.1109/ICASSP.2011.5947337>
- [58] P. Cui, X. Wang, J. Pei, and W. Zhu. 2019. A Survey on Network Embedding. *IEEE Transactions on Knowledge and Data Engineering* 31, 6 (2019), 833–852. <https://doi.org/10.1109/TKDE.2018.2849727>
- [59] J. Cukier. 2013. Making the Game of Thrones visualization. <http://www.jeromecukier.net/blog/2013/05/13/making-the-game-of-thrones-visualization/>. (05 2013).
- [60] L. da Fontoura Costa, F. A. Rodrigues, G. Travieso, and P. R. Villas Boas. 2007. Characterization of complex networks: A survey of measurements. *Advances in Physics* 56, 1 (2007), 167–242. <https://doi.org/10.1080/00018730601170527>
- [61] N. Dekker, T. Kuhn, and M. van Erp. 2019. Evaluating named entity recognition tools for extracting social networks from novels. *PeerJ Computer Science* 5 (2019), e189. <https://doi.org/10.7717/peerj-cs.189>

- [62] L. A. Deleris, F. Bonin, E. Daly, S. Deparis, Y. Hou, C. Jochim, Y. Lassoued, and K. Levacher. 2018. Know Who Your Friends Are: Understanding Social Connections from Unstructured Text. In *NAACL-HLT Demonstrations*. 76–80. <http://aclweb.org/anthology/N18-5016>
- [63] L. Ding and A. Yilmaz. 2010. Learning Relations Among Movie Characters: A Social Networks Perspective. In *European Conference on Computer Vision (Lecture Notes in Computer Science)*, Vol. 6314. 410–423. https://doi.org/10.1007/978-3-642-15561-1_30
- [64] L. Ding and A. Yilmaz. 2011. Inferring social relations from visual concepts. In *International Conference on Computer Vision*. 699–706. <https://doi.org/10.1109/ICCV.2011.6126306>
- [65] T. T. H. Do, Q. H. B. Tran, and Q. D. Tran. 2018. Movie indexing and summarization using social network techniques. *Vietnam Journal of Computer Science* 5, 2 (2018), 157–164. <https://doi.org/10.1007/s40595-018-0111-2>
- [66] P. Doreian. 2017. Reflections on Studying Signed Networks. *Journal of Interdisciplinary Methodologies and Issues in Science* 2 (2017), 2.1–2.14. <https://doi.org/10.18713/JIMIS-170117-2-1>
- [67] N. Dugué, V. Labatut, and A. Perez. 2015. A community role approach to assess social capitalists visibility in the Twitter network. *Social Network Analysis and Mining* 5 (2015), 26. <https://doi.org/10.1007/s13278-015-0266-0>
- [68] M. Edwards, L. Mitchell, J. Tuke, and M. Roughan. 2018. The one comparing narrative social network extraction techniques. *arXiv cs.SI* (2018), 1811.01467. <https://arxiv.org/abs/1811.01467>
- [69] M. Elsner. 2012. Character-based Kernels for Novelistic Plot Structure. In *13th Conference of the European Chapter of the Association for Computational Linguistics*. 634–644. <http://www.aclweb.org/anthology/E12-1065>
- [70] D. K. Elson. 2012. *Modeling Narrative Discourse*. Ph.D. Dissertation. http://www.cs.columbia.edu/~delson/pubs/Modeling-Narrative-Discourse_Elson_R4.pdf
- [71] D. K. Elson, N. Dames, and K. R. McKeown. 2010. Extracting social networks from literary fiction. In *48th annual meeting of the association for computational linguistics*. 138–147. <http://dl.acm.org/citation.cfm?id=1858696>
- [72] D. K. Elson and K. R. McKeown. 2010. Automatic attribution of quoted speech in literary narrative. In *24th AAAI Conference on Artificial Intelligence*. 1013–1019. <https://dl.acm.org/citation.cfm?id=2898607.2898769>
- [73] P. Ercolessi. 2013. *Extraction multimodale de la structure narrative des épisodes de séries télévisées*. PhD Thesis. Université de Toulouse. <http://www.theses.fr/2013TOU30131>
- [74] M. Everingham, J. Sivic, and A. Zisserman. 2009. Taking the bite out of automated naming of characters in TV video. *Image and Vision Computing* 27, 5 (2009), 545–559. <https://dl.acm.org/citation.cfm?id=1519685>
- [75] M. Falk. 2016. Making Connections: Network Analysis, the Bildungsroman and the World of The Absentee. *Journal of Language, Literature and Culture* 63, 2-3 (2016), 107–122. <https://doi.org/10.1080/20512856.2016.1244909>
- [76] F. Fischer. 2018. Network Analysis of the 9th Season of "Curb your Enthusiasm". <https://weltliteratur.net/curb-your-enthusiasm-season-9-network-analysis/>. (2018).
- [77] F. Fischer, G. Dazord, M. Göbel, C. Kittel, and P. Trilcke. 2017. Le drame comme réseau de relations : Une application de l'analyse automatisée pour l'histoire littéraire du théâtre. *Revue d'Historiographie du Théâtre* 4 (2017), 1–9. <https://sht.asso.fr/le-drame-comme-reseau-de-relations-une-application-de-lanalyse-automatisee-pour-lhistoire-litteraire-du-theatre/>
- [78] F. Fischer, M. Göbel, D. Kampkaspar, C. Kittel, and P. Trilcke. 2017. Network Dynamics, Plot Analysis: Approaching the Progressive Structuration of Literary Texts. In *Digital Humanities*. <https://dh2017.adho.org/abstracts/071/071.pdf>
- [79] F. Fischer, M. Göbel, D. Kampkaspar, and P. Trilcke. 2015. Digital Network Analysis of Dramatic Texts. In *Digital Humanities*. http://dh2015.org/abstracts/xml/FISCHER_Frank_Digital_Network_Analysis_of_Dramati/FISCHER_Frank_Digital_Network_Analysis_of_Dramatic_Text.html
- [80] F. Fischer, A. Schultz, C. Kittel, E. Beshero-Bondar, S. Martus, P. Trilcke, J. Wolf, I. Börner, A. Hechtel, D. Skorinkin, T. Orlova, C. Milling, and C. Ivanovic. 2018. Brecht Beats Shakespeare! – A Card-Game Introduction to the Network Analysis of European Drama. <https://github.com/lehkost/dramenquartett>. (2018). <https://doi.org/10.6084/m9.figshare.6667424.v1>
- [81] S. Fortunato. 2010. Community detection in graphs. *Physics Reports* 486, 3-5 (2010), 75–174. <https://doi.org/10.1016/j.physrep.2009.11.002>
- [82] L. C. Freeman. 1977. A Set of Measures of Centrality Based on Betweenness. *Sociometry* 40, 1 (1977), 35–41. <https://doi.org/10.2307/3033543>
- [83] L. C. Freeman. 1978. Centrality in Social Networks I: Conceptual Clarification. *Social Networks* 1, 3 (1978), 215–239. [https://doi.org/10.1016/0378-8733\(78\)90021-7](https://doi.org/10.1016/0378-8733(78)90021-7)
- [84] E. Gabasova. 2015. The Star Wars social network. <http://evelinag.com/blog/2015/12-15-star-wars-social-network/index.html>. (12 2015).
- [85] E. Gabasova. 2016. Star Wars social networks: The Force Awakens. <http://evelinag.com/blog/2016/01-25-social-network-force-awakens/index.html>. (01 2016).
- [86] N. P. Garg, S. Favre, H. Salamin, D. Hakkani Tür, and A. Vinciarelli. 2008. Role Recognition for Meeting Participants: an Approach Based on Lexical Information and Social Network Analysis. In *16th ACM international conference on*

- Multimedia*. 693–696. <https://doi.org/10.1145/1459359.1459462>
- [87] S. Ghannay, A. Caubriere, Y. Estève, N. Camelin, E. Simonnet, A. Laurent, and E. Morin. 2018. End-To-End Named Entity And Semantic Concept Extraction From Speech. In *IEEE Spoken Language Technology Workshop*. <https://doi.org/10.1109/slt.2018.8639513>
- [88] M. Girvan and M. E. J. Newman. 2002. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences* 99, 12 (2002), 7821–7826. <https://doi.org/10.1073/pnas.122653799>
- [89] S. Glander. 2017. Network analysis of Game of Thrones family ties. https://shiring.github.io/networks/2017/05/15/got_final. (05 2017).
- [90] K. Glass and S. Bangay. 2006. Hierarchical rule generalisation for speaker identification in fiction books. In *Annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*. 31–40. <https://doi.org/10.1145/1216262.1216266>
- [91] K. Glass and S. Bangay. 2007. A naive salience-based method for speaker identification in fiction books. In *18th Annual Symposium of the Pattern Recognition Association of South Africa*. 1312–1320. <http://www.prasa.org/proceedings/2007/prasa07-05.pdf>
- [92] P. M. Gleiser. 2007. How to become a superhero. *Journal of Statistical Mechanics* 2007, 09 (2007), P09020. <https://doi.org/10.1088/1742-5468/2007/09/P09020>
- [93] H.-N. Goh, L.-K. Soon, and S.-C. Haw. 2012. Automatic identification of protagonist in fairy tales using verb. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining (Lecture Notes in Computer Science)*, Vol. 7302. 395–406. https://doi.org/10.1007/978-3-642-30220-6_33
- [94] P. J. Gorinski and M. Lapata. 2015. Movie Script Summarization as Graph-based Scene Extraction. In *Annual Conference of the North American Chapter of the ACL*. 1066–1076. <http://www.aclweb.org/anthology/N15-1113>
- [95] A. Goyal, V. Gupta, and M. Kumar. 2018. Recent Named Entity Recognition and Classification techniques: A systematic review. *Computer Science Review* 29 (2018), 21–43. <https://doi.org/10.1016/j.cosrev.2018.06.001>
- [96] M. Grandjean. 2015. Network visualization: mapping Shakespeare’s tragedies. <http://www.martingrandjean.ch/network-visualization-shakespeare/>. (12 2015).
- [97] M. Grandjean. 2017. Comparing the Relational Structure of the Gospels: Network Analysis as a Tool for Biblical Narratology. *Digital Biblical Studies* 2 (2017), in press. <http://www.martingrandjean.ch/communications/>
- [98] S. Grayson, K. Wade, G. Meaney, and D. Greene. 2016. The Sense and Sensibility of Different Sliding Windows in Constructing Co-occurrence Networks from Literature. In *International Workshop on Computational History and Data-Driven Humanities*. 65–77. https://doi.org/10.1007/978-3-319-46224-0_7
- [99] A. Grener, M. Luczak-Roesch, E. Fenton, and T. Goldfinch. 2017. *Towards a Computational Literary Science: A Computational Approach to Dickens’ Dynamic Character Networks*. Technical Report. Zenodo. <https://doi.org/10.5281/zenodo.259499>
- [100] R. Guimerà and L. A. N. Amaral. 2005. Functional cartography of complex metabolic networks. *Nature* 433 (2005), 895–900. <https://doi.org/10.1038/nature03288>
- [101] F. Harary. 1969. *Graph Theory*. Addison-Wesley. <http://www.dtic.mil/dtic/tr/fulltext/u2/705364.pdf>
- [102] H. He, D. Barbosa, and G. Kondrak. 2013. Identification of speakers in novels. In *51st Annual Meeting of the Association for Computational Linguistics*. 1312–1320. <http://www.aclweb.org/anthology/P13-1129>
- [103] J. He, Y. Xie, X. Luan, L. Zhang, and X. Zhang. 2018. SRN: The Movie Character Relationship Analysis via Social Network. In *International Conference on Multimedia Modeling (Lecture Notes in Computer Science)*, Vol. 10705. Springer, 289–301. https://doi.org/10.1007/978-3-319-73600-6_25
- [104] F. Heider. 1946. Attitudes and cognitive organization. *Journal of Psychology* 21, 1 (1946), 107–112. <https://doi.org/10.1080/00223980.1946.9917275>
- [105] L. Hettlinger, M. Becker, I. Reger, F. Jannidis, and A. Hotho. 2015. Genre Classification on German Novels. In *26th International Workshop on Database and Expert Systems Applications*. <https://doi.org/10.1109/DEXA.2015.62>
- [106] R. Hickman. 2016. A Plot of Networks (Part 1). <https://roberthickman.website/2016/12/01/a-plot-of-networks-part-1/>. (12 2016).
- [107] H. N. Ho, C. Rigaud, J.-C. Burie, and J.-M. Ogier. 2013. Detecting Recurring Deformable Objects: An Approximate Graph Matching Method for Detecting Characters in Comics Books. In *International Workshop on Graphics Recognition (Lecture Notes in Computer Science)*, Vol. 8746. 122–134. https://doi.org/10.1007/978-3-662-44854-0_10
- [108] A. J. Holanda, M. Matias, S. M. S. P. Ferreira, G. M. L. Benevides, and O. Kinouchi. 2017. Character Networks and Book Genre Classification. *arXiv cs.SI* (2017), 1704.08197. <https://arxiv.org/abs/1704.08197>
- [109] P. Holme and J. Saramäki. 2012. Temporal networks. *Physics Reports* 519, 3 (2012), 97–125. <https://doi.org/10.1016/j.physrep.2012.03.001>
- [110] S. Hutchinson, V. Datla, and M. M. Louwerse. 2012. Social networks are encoded in language. In *34th annual conference of the cognitive science society*. 491–496. <https://mindmodeling.org/cogsci2012/papers/0096/paper0096.pdf>

- [111] iuhfd. 2018. Visualisation of words spoken between Romeo and Juliet characters. https://www.reddit.com/r/DataArt/comments/8iwax2/visualisation_of_words_spoken_between_romeo_and/. (05 2018).
- [112] M. Iyyer, A. Guha, S. Chaturvedi, J. Boyd-Graber, and H. Daumé. 2016. Feuding families and former friends: Unsupervised learning for dynamic fictional relationships. In *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 1534–1544. <https://doi.org/10.18653/v1/N16-1180>
- [113] F. Jannidis, I. Reger, M. Krug, L. Weimer, L. Macharowsky, and F. Puppe. 2016. Comparison of Methods for the Identification of Main Characters in German Novels. In *Digital Humanities*. 578–582. <http://dh2016.adho.org/abstracts/297>
- [114] M. Janosov. 2017. Network Science Predicts Who Dies Next in Game of Thrones. <https://networkdatascience.ceu.edu/article/2017-07-08/network-science-predicts-who-dies-next-game-thrones>. (07 2017).
- [115] P. A. Jayannavar, A. Agarwal, M. Ju, and O. Rambo. 2015. Validating literary theories using automatic social network extraction. In *4th Workshop on Computational Linguistics for Literature*. 32–41. <http://anthology.aclweb.org/W/W15/W15-07.pdf#page=46>
- [116] J. Jovanovic, E. Bagheri, J. Cuzzola, D. Gasevic, Z. Jeremic, and R. Bashash. 2014. Automated semantic tagging of textual content. *IT Professional* 16, 6 (2014), 38–46. <https://doi.org/10.1109/MITP.2014.85>
- [117] B. Jung, T. Kwak, J. Song, and Y. Lee. 2004. Narrative abstraction model for story-oriented video. In *12th annual ACM international conference on Multimedia*. 828–835. <https://doi.org/10.1145/1027527.1027720>
- [118] J. J. Jung, E. You, and S.-B. Park. 2013. Emotion-based character clustering for managing story-based contents: a cinemetric analysis. *Multimedia Tools and Applications* 65, 1 (2013), 29–45. <https://doi.org/10.1007/s11042-012-1133-x>
- [119] F. Karsdorp, P. van Kranenburg, T. Meder, and A. van den Bosch. 2012. Casting a spell: Identification and ranking of actors in folktales. In *2nd Workshop on Annotation of Corpora for Research in the Humanities*. https://pure.knaw.nl/ws/files/481270/karsdorp_et_al2012b.pdf
- [120] A. Kazantseva and S. Szpakowicz. 2010. Summarizing short stories. *Computational Linguistics* 36, 1 (2010), 71–109. <https://doi.org/10.1162/coli.2010.36.1.36102>
- [121] R. Kenna and P. Mac Carron. 2016. Maths meets myths: Network investigations of ancient narratives. In *International Conference on Computer Simulation in Physics and Beyond (Journal of Physics: Conference Series)*, Vol. 681. 012002. <https://doi.org/10.1088/1742-6596/681/1/012002>
- [122] R. Kenna and P. MacCarron. 2017. A Networks Approach to Mythological Epics. In *Maths Meets Myths: Quantitative Approaches to Ancient Narratives*. Springer, 21–43. https://doi.org/10.1007/978-3-319-39445-9_3
- [123] B. Ker, C. W. Xian, and D. Chong. 2018. A who’s who guide to the Marvel Cinematic Universe. <https://graphics.straitstimes.com/STI/STIMEDIA/Interactives/2018/04/marvel-cinematic-universe-whos-who-interactive/index.html>. (04 2018).
- [124] H. J. Kim and J. M. Kim. 2005. Cyclic topology in complex networks. *Physical Review E* 72, 3 (2005), 036109. <https://doi.org/10.1103/PhysRevE.72.036109>
- [125] J. Kleinberg. 1999. Authoritative sources in a hyperlinked environment. *Journal of the Association for Computing Machinery* 46, 5 (1999), 604–632. <https://doi.org/10.1145/324133.324140>
- [126] D. E. Knuth. 1993. *The Stanford GraphBase: a platform for combinatorial computing*. Vol. 37. Addison-Wesley Reading. <http://www-cs-faculty.stanford.edu/~knuth/sgb.html>
- [127] D. Kokkinakis and M. Malm. 2011. Character Profiling in 19th Century Fiction. In *Language Technologies for Digital Humanities and Cultural Heritage Workshop*. 70–77. <http://aclweb.org/anthology/W11-4111>
- [128] D. Koschützki, K. A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski. 2005. Centrality Indices. In *Network Analysis - Methodological Foundations*. Lecture Notes in Computer Science, Vol. 3418. Springer, 16–61. https://doi.org/10.1007/978-3-540-31955-9_3
- [129] V. Krishnan and J. Eisenstein. 2015. "You're Mr. Lebowsky, I'm the Dude": Inducing Address Term Formality in Signed Social Networks. In *Annual Conference of the North American Chapter of the ACL*. 1616–1626. <http://www.aclweb.org/anthology/N15-1185>
- [130] M. Krug, F. Puppe, F. Jannidis, L. Macharowsky, I. Reger, and L. Weimer. 2015. Rule-based Coreference Resolution in German Historic Novels. In *4th Workshop on Computational Linguistics for Literature*. 98–104. <http://www.aclweb.org/anthology/W15-0711>
- [131] M. Krug, I. Reger, F. Jannidis, L. Weimer, N. Madarász, and F. Puppe. 2017. Overcoming Data Sparsity for Relation Detection in German Novels. In *Digital Humanities*. <https://dh2017.adho.org/abstracts/340/340.pdf>
- [132] D. Kruger, M. Fisher, and I. Jobling. 2003. Proper and dark heroes as DADS and CADS. *Human Nature* 14, 3 (2003), 305–317. <https://doi.org/10.1007/s12110-003-1008-y>
- [133] A. Kumar, A. Kaur, and M. Kumar. 2018. Face detection techniques: a review. *Artificial Intelligence Review* in press (2018). <https://doi.org/10.1007/s10462-018-9650-2>
- [134] J. Kunegis. 2019. Network of Characters in Harry Potter. <https://networkscience.wordpress.com/2019/06/07/network-of-characters-in-harry-potter/>. (06 2019).

- [135] H.-C. Kwon and K.-H. Shim. 2017. An Improved Method of Character Network Analysis for Literary Criticism: A Case Study of Hamlet. *International Journal of Contents* 13, 3 (2017), 43–48. <https://doi.org/10.5392/IJoC.2017.13.3.043>
- [136] D. Kydros and A. Anastasiadis. 2015. Social network analysis in literature. The case of The Great Eastern by A. Embirikos. In *5th European Congress of Modern Greek Studies of the European Society of Modern Greek Studies*. 681–702. https://www.researchgate.net/profile/Dimitrios_Kydros/publication/273357563_Social_network_analysis_in_literature_The_case_of_The_Great_Eastern_by_A_Embirikos/links/54ff172f0cf2672e22419bd6.pdf
- [137] D. Kydros, P. Notopoulos, and G. Exarchos. 2015. Homer’s Iliad - a social network analytic approach. *International Journal of Humanities and Arts Computing* 9 (2015), 115–132. <https://doi.org/10.3366/ijhac.2015.0141>
- [138] V. Labatut. 2014. Étude de l’omniprésence des propriétés petit-monde et sans-échelle. In *5ème Conférence sur les modèles et l’analyse de réseaux : approches mathématiques et informatiques*. <http://lipn.univ-paris13.fr/marami/MARAMI/pub2014.html>
- [139] J. Lee and T.-S. Wong. 2016. Conversational Network in the Chinese Buddhist Canon. *Open Linguistics* 2, 1 (2016), 427–436. <https://doi.org/10.1515/opli-2016-0022>
- [140] J. Lee and C. Y. Yeung. 2012. Extracting networks of people and places from literary texts. In *26th Pacific Asia Conference on Language, Information and Computation*. 209–218. <https://aclanthology.info/papers/Y12-1022/y12-1022>
- [141] O.-J. Lee and J. J. Jung. 2016. Affective Character Network for Understanding Plots of Narrative Contents. In *Affective Computing and Context Awareness in Ambient Intelligence*. <http://ceur-ws.org/Vol-1794/afcai16-paper3.pdf>
- [142] O.-J. Lee and J. J. Jung. 2019. Character Network Embedding-based Plot Structure Discovery in Narrative Multimedia. In *9th International Conference on Web Intelligence, Mining and Semantics*. 15. <https://doi.org/10.1145/3326467.3326485>
- [143] O.-J. Lee and J. J. Jung. 2019. Integrating character networks for extracting narratives from multimodal data. *Information Processing & Management* 56, 5 (2019), 1894–1923. <https://doi.org/10.1016/j.ipm.2019.02.005>
- [144] O.-J. Lee and J. J. Jung. 2019. Modeling affective character network for story analytics. *Future Generation Computer Systems* 92 (2019), 458–478. <https://doi.org/10.1016/j.future.2018.01.030>
- [145] J. Li, D. Yang, and P. Lv. 2018. Visualize classic play’s composing patterns: a weighted motif mining framework. *Multimedia Tools and Applications* in press (2018). <https://doi.org/10.1007/s11042-018-6405-7>
- [146] J.-Y. Li, L.-W. Kang, C.-M. Tsai, and C.-W. Lin. 2015. Learning-based movie summarization via role-community analysis and feature fusion. In *17th IEEE International Workshop on Multimedia Signal Processing*. <https://doi.org/10.1109/MMSP.2015.7340794>
- [147] Y. Li, S. Narayanan, and C. C. J. Kuo. 2004. Content-based movie analysis and indexing based on audiovisual cues. *IEEE Transactions on Circuits and Systems for Video Technology* 14, 8 (2004), 1073–1085. <https://doi.org/10.1109/TCSVT.2004.831968>
- [148] C. Liang, Y/ Zhang, J. Cheng, C.-S. Xu, and H-Q. Lu. 2009. A novel role-based movie scene segmentation method. In *Pacific-Rim Conference on Multimedia (Lecture Notes in Computer Science)*, Vol. 5879. 917–922. https://doi.org/10.1007/978-3-642-10467-1_82
- [149] Ivan Lillo, Alvaro Soto, and Juan Carlos Niebles. 2014. Discriminative Hierarchical Modeling of Spatio-temporally Composable Human Activities. In *IEEE Conference on Computer Vision and Pattern Recognition*. 812–819. <https://doi.org/10.1109/CVPR.2014.109>
- [150] B. Lind. 2012. Lessons on exponential random graph modeling from Grey’s Anatomy hook-ups. <http://badhessian.org/2012/09/lessons-on-exponential-random-graph-modeling-from-greys-anatomy-hook-ups/>. (04 2012).
- [151] D. Liu and L. Albergante. 2017. Balance of thrones: a network study on ‘Game of Thrones’. *arXiv cs.SI* (2017), 1707.05213. <https://arxiv.org/abs/1707.05213>
- [152] R. D. Luce and A. D. Perry. 1949. A method of matrix analysis of group structure. *Psychometrika* 14, 2 (1949), 95–116. <https://doi.org/10.1007/BF02289146>
- [153] J. Lv, B. Wu, L. Zhou, and H. Wang. 2018. StoryRoleNet: Social Network Construction of Role Relationship in Video. *IEEE Access* 6 (2018), 25958–25969. <https://doi.org/10.1109/ACCESS.2018.2832087>
- [154] P. Mac Carron and R. Kenna. 2012. Universal Properties of Mythological Networks. *Europhysics Letters* 99, 2 (2012), 28002. <https://doi.org/10.1209/0295-5075/99/28002>
- [155] P. Mac Carron and R. Kenna. 2013. Network analysis of the Íslendinga sögur - the Sagas of Icelanders. *European Physical Journal B* 86 (2013), 407. <https://doi.org/10.1140/epjb/e2013-40583-3>
- [156] P. Mac Carron and R. Kenna. 2013. Viking sagas: Six degrees of Icelandic separation Social networks from the Viking era. *Significance* 10, 6 (2013), 12–17. <https://doi.org/10.1111/j.1740-9713.2013.00704.x>
- [157] P. Mac Carron and R. Kenna. 2014. A quantitative approach to comparative mythology. *Cosmos* 14 (2014), 103–117. <https://ora.ox.ac.uk/objects/uuid:43275d59-4db2-49db-b26a-96f21378cfa8>
- [158] P. MacCarron. 2014. *A Network Theoretic Approach to Comparative Mythology*. PhD Thesis. <https://curve.coventry.ac.uk/open/file/9bfc043d-497e-4217-82ab-e19963b53290/1/maccarroncomb.pdf>
- [159] A. Makazhanov, D. Barbosa, and G. Kondrak. 2014. Extracting Family Relationship Networks from Novels. *arXiv cs.CL* (2014), 1405.0603. <https://arxiv.org/abs/1405.0603>

- [160] C. Makris and P. Vikatos. 2016. Community Detection of Screenplay Characters. In *12th IFIP International Conference on Artificial Intelligence Applications and Innovations (IFIP Advances in Information and Communication Technology)*, Vol. 475. 463–470. https://doi.org/10.1007/978-3-319-44944-9_40
- [161] N. Mamede and P. Chaleira. 2004. Character Identification in Children Stories. In *International Conference on Natural Language Processing in Spain (Lecture Notes in Computer Science)*, Vol. 3230. 82–90. https://doi.org/10.1007/978-3-540-30228-5_8
- [162] S. A. Marvel, J. Kleinberg, R. D. Kleinberg, and S. H. Strogatz. 2011. Continuous-time model of structural balance. *Proceedings of the National Academy of Sciences* 108, 5 (2011), 1771–1776. <https://doi.org/10.1073/pnas.1013213108>
- [163] V. H. Masías, P. Baldwin, S. Laengle, A. Vargas, and F. A. Crespo. 2017. Exploring the prominence of Romeo and Juliet’s characters using weighted centrality measures. *Digital Scholarship in the Humanities* 32, 4 (2017), 837–858. <https://doi.org/10.1093/lc/fqw029>
- [164] D. McNary. 2019. 2018 Worldwide Box Office Hits Record as Disney Dominates. <https://variety.com/2019/film/news/box-office-record-disney-dominates-1203098075/>. (01 2019).
- [165] S. Min and J. Park. 2016. Mapping Out Narrative Structures and Dynamics Using Networks and Textual Information. *arXiv cs.CL* (2016), 1604.03029. <http://arxiv.org/abs/1604.03029>
- [166] S. Min and J. Park. 2016. Narrative as a complex network: A study of Victor Hugo’s *Les Misérables*. In *HCI Korea*. 100–107. <https://doi.org/10.17210/hcik.2016.01.100>
- [167] S. Min and J. Park. 2016. Network Science and Narratives: Basic Model and Application to Victor Hugo’s *Les Misérables*. In *7th Workshop on Complex Networks CompleNet (Studies in Computational Intelligence)*, Vol. 644. 257–265. https://doi.org/10.1007/978-3-319-30569-1_19
- [168] P. J. Miranda, M. S. Baptista, and S. E. de S. Pinto. 2013. Analysis of communities in a mythological social network. *arXiv physics.soc-ph* (2013), 1306.2537. <http://arxiv.org/abs/1306.2537>
- [169] M. M. Moattar, M. H. and Homayounpour. 2012. A review on speaker diarization systems and approaches. *Speech Communication* 54, 10 (2012), 1065–1103. <https://doi.org/10.1016/j.specom.2012.05.002>
- [170] F. Moretti. 2011. Network Theory, Plot Analysis. *Stanford Literary Lab* 2 (2011). <http://litlab.stanford.edu/LiteraryLabPamphlet2.pdf>
- [171] Y. Mourchid, B. Renoust, H. Cherifi, and M. El Hassouni. 2018. Multilayer Network Model of Movie Script. In *7th International Conference on Complex Networks and Their Applications (Studies in Computational Intelligence)*, Vol. 812. Springer, 782–796. https://doi.org/10.1007/978-3-030-05411-3_62
- [172] S. Muhuri, S. Chakraborty, and S. N. Chakraborty. 2018. Extracting Social Network and Character Categorization From Bengali Literature. *IEEE Transactions on Computational Social Systems* 5, 2 (2018), 371–381. <https://doi.org/10.1109/TCSS.2018.2798699>
- [173] P. Mutton. 2004. Inferring and visualizing social networks on internet relay chat. In *8th International Conference on Information Visualisation*. 35–43. <https://doi.org/10.1109/IV.2004.1320122>
- [174] D. Nadeau and S. Sekine. 2007. A survey of named entity recognition and classification. *Linguisticae Investigationes* 30, 1 (2007), 3–26. <https://doi.org/10.1075/li.30.1.03nad>
- [175] E. T. Nalisnick and H. S. Baird. 2013. Character-to-Character Sentiment Analysis in Shakespeare’s Plays. In *51st Annual Meeting of the Association for Computational Linguistics*. <http://www.aclweb.org/anthology/P13-2085>
- [176] E. T. Nalisnick and H. S. Baird. 2013. Extracting Sentiment Networks from Shakespeare’s Plays. In *12th International Conference on Document Analysis and Recognition*. 758–762. <https://doi.org/10.1109/ICDAR.2013.155>
- [177] C.-J. Nan, K.-M. Kim, and B.-T. Zhang. 2015. Social Network Analysis of TV Drama Characters Via Deep Concept Hierarchies. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. 831–836. <https://doi.org/10.1145/2808797.2809306>
- [178] M. E. J. Newman. 2002. Assortative Mixing in Networks. *Physical Review Letters* 89, 20 (2002), 208701. <https://doi.org/10.1103/PhysRevLett.89.208701>
- [179] M. E. J. Newman. 2003. Mixing patterns in networks. *Physical Review E* 67 (2003), 026126. <https://doi.org/10.1103/PhysRevE.67.026126>
- [180] M. E. J. Newman. 2003. The structure and function of complex networks. *SIAM Rev.* 45 (2003), 167–256. <https://doi.org/10.1137/S003614450342480>
- [181] M. E. J. Newman and M. Girvan. 2004. Finding and evaluating community structure in networks. *Physical Review E* 69, 2 (2004), 026113. <https://doi.org/10.1103/PhysRevE.69.026113>
- [182] M. Nijila and M. T. Kala. 2018. Extraction of Relationship Between Characters in Narrative Summaries. In *International Conference on Emerging Trends and Innovations In Engineering And Technological Research*. 1–5. <https://doi.org/10.1109/ICETIETR.2018.8529125>
- [183] notrudedude. 2018. Who kissed whom in The Office. <https://i.imgur.com/oiiyKlz.png>. (06 2018).
- [184] D. Oelke, D. Kokkinakis, and D. A. Keim. 2013. Fingerprint Matrices: Uncovering the dynamics of social networks in prose literature. *Computer Graphics Forum* 32, 3pt4 (jun 2013), 371–380. <https://doi.org/10.1111/cgf.12124>

- [185] D. Oelke, D. Kokkinakis, and M. Malm. 2012. Advanced Visual Analytics Methods for Literature Analysis. In *6th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*. 35–44. <https://dl.acm.org/citation.cfm?id=2390364>
- [186] T. O’Keefe, S. Pareti, J. R. Curran, I. Koprinska, and M. Honnibal. 2012. A sequence labelling approach to quote attribution. In *Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 790–799. <https://dl.acm.org/citation.cfm?id=2391033>
- [187] É. Papégnies, V. Labatut, R. Dufour, and G. Linarès. 2019. Conversational Networks for Automatic Online Moderation. *IEEE Transactions on Computational Social Systems* 6, 1 (2019), 38–55. <https://doi.org/10.1109/TCSS.2018.2887240>
- [188] G.-M. Park, S.-H. Kim, and H.-G. Cho. 2013. Structural analysis on social network constructed from characters in literature texts. *Journal of Computers* 8, 9 (2013), 2442–2447. <http://www.jcomputers.us/vol8/jcp0809-38.pdf>
- [189] G.-M. Park, S.-H. Kim, H.-R. Hwang, and H.-G. Cho. 2013. Complex System Analysis of Social Networks Extracted from Literary Fictions. *International Journal of Machine Learning and Computing* 1 (2013), 107–111. <https://doi.org/10.7763/IJMLC.2013.V3.282>
- [190] S.-B. Park, Y.-W. Kim, M. N. Uddin, and G.-S. Jo. 2009. Character-Net: Character Network Analysis from Video. In *IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies*. 305–308. <https://doi.org/10.1109/WI-IAT.2009.54>
- [191] S.-B. Park, K.-J. Oh, and G.-S. Jo. 2012. Social network analysis in a movie using Character-Net. *Multimedia Tools and Applications* 59, 2 (jan 2012), 601–627. <https://doi.org/10.1007/s11042-011-0725-1>
- [192] J. Pennington, R. Socher, and C. Manning. 2014. Glove: Global vectors for word representation. In *Conference on Empirical Methods in Natural Language Processing*. 1532–1543. <https://doi.org/10.3115/v1/D14-1162>
- [193] M. Pettit. 2016. Who’s at the Center of the Star Trek Universe? <http://www.datascopeanalytics.com/blog/star-trek-viz/>. (07 2016).
- [194] Massimo Poesio, Roland Stuckardt, and Yannick Versley. 2016. *Anaphora Resolution: Algorithms, Resources, and Applications*. Springer. <https://doi.org/10.1007/978-3-662-47909-4>
- [195] F. Pope, R. A. Shirvani, Mugizi R. Rwebangira, M. Chouikha, A. Taylor, A. Alarcon Ramirez, and A. Torfi. 2015. Automatic Detection of Small Groups of Persons, Influential Members, Relations and Hierarchy in Written Conversations Using Fuzzy Logic. In *International Conference on Data Mining*. 155–161. <https://arxiv.org/abs/1610.01720v1>
- [196] R. Poppe. 2010. A survey on vision-based human action recognition. *Image and Vision Computing* 28, 6 (2010), 976–990. <https://doi.org/10.1016/j.imavis.2009.11.014>
- [197] S. D. Prado, S. R. Dahmen, A. L. C. Bazzan, P. Mac Carron, and R. Kenna. 2016. Temporal Network Analysis of Literary Texts. *Advances in Complex Systems* 19, 3 (2016), 1650005. <https://doi.org/10.1142/S0219525916500053>
- [198] C. Pudney. 2014. The Simpsons Social Network (Season 1). (2014).
- [199] M. A. Ribeiro, R. A. Vosgerau, M. L. Pereira Andruchiw, and S. E. de Souza Pinto. 2016. The complex social network from The Lord of The Rings. *Revista Brasileira de Ensino de Física* (2016). https://www.researchgate.net/profile/Sandro_Pinto/publication/286449748_The_complex_social_network_from_The_Lord_of_The_Rings/links/5669dfbe08ae62b05f026df9.pdf
- [200] B. Rieck and H. Leitte. 2016. ‘Shall I compare thee to a network?’: Visualizing the Topological Structure of Shakespeare’s Plays. In *1st Workshop on Visualization for the Digital Humanities*. <https://doi.org/10.11588/heidok.00023477>
- [201] C. Rigaud, T. N. Le, J.-C. Burie, J.-M. Ogier, M. Iwata, E. Imazu, and K. Kise. 2015. Speech balloon and speaker association for comics and manga understanding. In *13th International Conference on Document Analysis and Recognition*. 351–355. <https://doi.org/10.1109/ICDAR.2015.7333782>
- [202] N. Rinehart. 2017. The Social Network of Harry Potter. <https://prezi.com/trmygkzvocu2/the-social-network-of-harry-potter/>. (2017).
- [203] D. Robinson. 2015. Analyzing networks of characters in ‘Love Actually’. <http://varianceexplained.org/r/love-actually-network/>. (12 2015). <http://varianceexplained.org/r/love-actually-network/>
- [204] D. Robson. 2016. Heroes and villains. *The Psychologist* 29, 8 (2016), 610–613. <https://thepsychologist.bps.org.uk/heroes-and-villains>
- [205] D. Robson. 2018. Our fiction addiction: Why humans need stories. <http://www.bbc.com/culture/story/20180503-our-fiction-addiction-why-humans-need-stories>. (05 2018).
- [206] Y. Rochat. 2014. *Character Networks and Centrality*. Ph.D. Dissertation. Université de Lausanne. http://infoscience.epfl.ch/record/203889/files/yrochat_thesis_infoscience.pdf
- [207] Y. Rochat. 2015. Character Network Analysis of Émile Zola’s Les Rougon-Macquart. In *Global Digital Humanities*. http://dh2015.org/abstracts/xml/ROCHAT_Yannick_Character_Network_Analysis_of__mil/ROCHAT_Yannick_Character_Network_Analysis_of__mile_Zola.html
- [208] Y. Rochat and F. Kaplan. 2014. Analyse de réseaux sur les personnages des Confessions de Jean-Jacques Rousseau. *Cahiers du numérique* 10, 3 (2014), 109–133. <https://doi.org/10.3166/LCN.10.3.109-133>

- [209] Y. Rochat and M. Triclot. 2017. Les réseaux de personnages de science-fiction : échantillons de lectures intermédiaires. *ReS Futuræ* 10 (2017). <https://doi.org/10.4000/resf.1183>
- [210] Y. Rui, T. S. Huang, and S. Mehrotra. 1999. Constructing table-of-content for videos. *Multimedia Systems* 7, 5 (1999), 359–368. <https://doi.org/10.1007/s005300050138>
- [211] D. Ruths. 2016. Why The Force Awakens is not just a remake of A New Hope. (03 2016).
- [212] M. Ryan and C. Littleton. 2017. TV Series Budgets Hit the Breaking Point as Costs Skyrocket in Peak TV Era. <https://variety.com/2017/tv/news/tv-series-budgets-costs-rising-peak-tv-1202570158/>. (09 2017).
- [213] J. Rydberg-Cox. 2011. Social Networks and the Language of Greek Tragedy. *Journal of the Chicago Colloquium on Digital Humanities and Computer Science* 1, 3 (2011). <https://letterpress.uchicago.edu/index.php/jdhcs/article/view/86>
- [214] G. A. Sack. 2012. Character Networks for Narrative Generation. In *8th Artificial Intelligence and Interactive Digital Entertainment Conference - Intelligent Narrative Technologies Workshop*. 38–43. <http://www.aaai.org/ocs/index.php/AIIDE/AIIDE12/paper/view/5550>
- [215] G. A. Sack. 2013. Character networks for narrative generation: Structural balance theory and the emergence of proto-narratives. In *Workshop on Computational Models of Narrative*. <http://drops.dagstuhl.de/opus/volltexte/2013/4161/pdf/p183-sack.pdf>
- [216] G. A. Sack. 2014. Character Networks for Narrative Generation: Structural Balance Theory and the Emergence of Proto-Narratives. In *Complexity and the Human Experience - Modeling Complexity in the Humanities and Social Sciences*. Pan Stanford Publishing - CRC Press, Chapter 4, 81–104. <https://doi.org/10.4032/9789814463270>
- [217] J. Sang, C. Liang, C.-S. Xu, and J. Cheng. 2011. Robust movie character identification and the sensitivity analysis. In *IEEE International Conference on Multimedia and Expo*. <https://doi.org/10.1109/ICME.2011.6011837>
- [218] J. Sang and C.-S. Xu. 2012. Robust Face-Name Graph Matching for Movie Character Identification. *IEEE Transactions on Multimedia* 14, 3 (2012), 586–596. <https://doi.org/10.1109/TMM.2012.2188784>
- [219] D. Schoch. 2015. Friends and Hypergraphs: The One With All The Networks. <http://mildlyscientific.schochastics.net/2015/03/03/friends-and-hypergraphs-one-with-a/>. (03 2015).
- [220] S. B. Seidman. 1983. Network structure and minimum degree. *Social Networks* 5, 3 (1983), 269–287. [https://doi.org/10.1016/0378-8733\(83\)90028-X](https://doi.org/10.1016/0378-8733(83)90028-X)
- [221] J. Seo, G.-M. Park, S.-H. Kim, and H.-G. Cho. 2013. Characteristic Analysis of Social Network Constructed from Literary Fiction. In *International Conference on Cyberworlds*. 147–150. <https://doi.org/10.1109/CW.2013.72>
- [222] J.-K. Seo, S.-H. Kim, H.-S. Tak, and H.-G. Cho. 2014. A structural analysis of literary fictions with social network framework. In *29th Annual ACM Symposium on Applied Computing*. 634–640. <https://doi.org/10.1145/2554850.2555049>
- [223] G. Simchoni. 2017. The one with friends. <http://giorasimchoni.com/2017/06/04/2017-06-04-the-one-with-friends/>. (06 2017).
- [224] D. Smith. 2017. Why do we tell stories? Hunter-gatherers shed light on the evolutionary roots of fiction. <https://theconversation.com/why-do-we-tell-stories-hunter-gatherers-shed-light-on-the-evolutionary-roots-of-fiction-88586>. (12 2017).
- [225] D. Smith, P. Schlaepfer, K. Major, M. Dyble, A. E. Page, J. Thompson, N. Chaudhary, G. D. Salali, R. Mace, L. Astete, M. Ngales, L. Vinicius, and A. Bamberg Migliano. 2017. Cooperation and the evolution of hunter-gatherer storytelling. *Nature Communications* 8 (2017), 1853. <https://doi.org/10.1038/s41467-017-02036-8>
- [226] K. Somandepalli, N. Kumar, T. Guha, and S. S. Narayanan. 2018. Unsupervised Discovery of Character Dictionaries in Animation Movies. *IEEE Transactions on Multimedia* 20, 3 (2018), 539–551. <https://doi.org/10.1109/TMM.2017.2745712>
- [227] sorted_hat. 2018. Graphing the character interactions from The Office, Seasons 1-7. https://www.reddit.com/r/dataisbeautiful/comments/8ikpap/oc_graphing_the_character_interactions_from_the/. (05 2018).
- [228] A. C. Sparavigna. 2013. On Social Networks in Plays and Novels. *International Journal of Sciences* 2, 10 (2013), 20–25. <https://doi.org/10.18483/ijSci.312>
- [229] S. Srivastava, S. Chaturvedi, and T. Mitchell. 2016. Inferring interpersonal relations in narrative summaries. In *30th AAAI Conference on Artificial Intelligence*. 2807–2813. <https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/viewPaper/12173>
- [230] J. Stavanja and M. Klemen. 2019. Predicting kills in Game of Thrones using network properties. *arXiv cs.SI* (2019), 1906.09468. <https://arxiv.org/abs/1906.09468>
- [231] J. Stiller and M. Hudson. 2005. Weak Links and Scene Cliques Within the Small World of Shakespeare. *Journal of Cultural and Evolutionary Psychology* 3, 1 (2005), 57–73. <https://doi.org/10.1556/JCEP.3.2005.1.4>
- [232] J. Stiller, D. Nettle, and R. I. M. Dunbar. 2003. The small world of Shakespeare’s plays. *Human Nature* 14, 4 (2003), 397–408. <https://doi.org/10.1007/s12110-003-1013-1>
- [233] S. Stoltzman. 2016. Seinfeld Characters - A Post About Nothing. <https://www.stoltzmaniac.com/seinfeld-characters-a-post-about-nothing/>. (12 2016).
- [234] M. Stommel, L. I. Merhej, and M. G. Müller. 2012. Segmentation-Free Detection of Comic Panels. In *International Conference on Computer Vision and Graphics (Lecture Notes in Computer Science)*, Vol. 7594. 633–640. <https://doi.org/>

[10.1007/978-3-642-33564-8_76](https://doi.org/10.1007/978-3-642-33564-8_76)

- [235] M. Stricker, O. Augereau, K. Kise, and M. Iwata. 2018. Facial Landmark Detection for Manga Images. *arXiv cs.CV* (2018), 1811.03214. <https://arxiv.org/abs/1811.03214>
- [236] S. Sudhahar and N. Cristianini. 2013. Automated Analysis of Narrative Content for Digital Humanities. *International Journal of Advanced Computer Science* 3, 9 (2013), 440–447. <https://doi.org/10.1.1.429.3639>
- [237] S. Sudhahar, G. De Fazio, R. Franzosi, and N. Cristianini. 2015. Network analysis of narrative content in large corpora. *Natural Language Engineering* 21 (2015), 81–112. <https://doi.org/10.1017/S1351324913000247>
- [238] C. Suen, L. Kuenzel, and S. Gil. 2013. Extraction and analysis of character interaction networks from plays and movies. In *Digital Humanities*. http://snap.stanford.edu/class/cs224w-2011/proj/laneyk_Finalwriteup_v1.pdf
- [239] Rhea Sukthanker, Soujanya Poria, Erik Cambria, and Ramkumar Thirunavukarasu. 2018. Anaphora and Coreference Resolution: A Review. *arXiv cs.CL* (2018), 1805.11824. <https://arxiv.org/abs/1805.11824>
- [240] W. Sun, J.-C. Burie, J.-M. Ogier, and K. Kise. 2013. Specific Comic Character Detection Using Local Feature Matching. In *12th International Conference on Document Analysis and Recognition*. 275–279. <https://doi.org/10.1109/ICDAR.2013.62>
- [241] K. Takayama, H. Johan, and T. Nishita. 2012. Face Detection and Face Recognition of Cartoon Characters Using Feature Extraction. In *IIEEJ Image Electronics and Visual Computing Workshop*. <http://www.iieej.org/trans/IEVC/IEVC2012/PDF/4B-1.pdf>
- [242] M. S. A. Tan, E. A. Ujum, and K. Ratnavelu. 2014. A character network study of two Sci-Fi TV series. In *Frontiers in Physics (AIP Conference Proceedings)*, Vol. 1588. AIP Publishing LLC, 246. <https://doi.org/10.1063/1.4866954>
- [243] M. Tandiwe Myambo. 2016. How reading fiction can help students understand the real world. <https://theconversation.com/how-reading-fiction-can-help-students-understand-the-real-world-52908>. (01 2016).
- [244] Q. D. Tran, D. Hwang, and J. J. Jung. 2015. Movie summarization using characters network analysis. In *7th International Conference on Computational Collective Intelligence (Lecture Notes in Computer Science)*, Vol. 9329. Springer, 390–399. https://doi.org/10.1007/978-3-319-24069-5_37
- [245] Q. D. Tran, D. Hwang, and J. J. Jung. 2017. Character-based indexing and browsing with movie ontology. *Journal of Intelligent & Fuzzy Systems* 32, 2 (2017), 1229–1240. <https://doi.org/10.3233/JIFS-169122>
- [246] Q. D. Tran, D. Hwang, O.-J. Lee, and J. E. Jung. 2017. Exploiting character networks for movie summarization. *Multimedia Tools and Applications* 76, 8 (2017), 10357–10369. <https://doi.org/10.1007/s11042-016-3633-6>
- [247] Q. D. Tran, D. Hwang, O.-J. Lee, and J. J. Jung. 2017. A novel method for extracting dynamic character network from movie. In *International Conference on Big Data Technologies and Applications (Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering)*, Vol. 194. 48–53. https://doi.org/10.1007/978-3-319-58967-1_6
- [248] Q. D. Tran and J. E. Jung. 2015. Cocharnet: Extracting social networks using character co-occurrence in movies. *Journal of Universal Computer Science* 21 (2015), 796–815. Issue 6. <https://doi.org/10.3217/jucs-021-06-0796>
- [249] Q. H. B. Tran, T. H. B. Nguyen, P. N. Tran, T. T. N. Tran, and Q. D. Tran. 2017. Story-Based Multimedia Analysis Using Social Network Technique. In *International Conference on Computer Science, Applied Mathematics and Applications (Advances in Intelligent Systems and Computing)*, Vol. 629. 95–106. https://doi.org/10.1007/978-3-319-61911-8_9
- [250] P. Trilcke, F. Fischer, M. Göbel, and D. Kampkaspar. 2016. Theatre Plays as ‘Small Worlds’? Network Data on the History and Typology of German Drama, 1730–1930. In *Digital Humanities*. 360. <http://dh2016.adho.org/abstracts/360>
- [251] M. Trovati and J. Brady. 2014. Towards an automated approach to extract and compare fictional networks: An initial evaluation. In *25th International Workshop on Database and Expert Systems Applications*. <https://doi.org/10.1109/DEXA.2014.58>
- [252] C.-M. Tsai, L.-W. Kang, C.-W. Lin, and W. Lin. 2013. Scene-Based Movie Summarization Via Role-Community Networks. *IEEE Transactions on Circuits and Systems for Video Technology* 23, 11 (2013), 1927–1940. <https://doi.org/10.1109/TCSVT.2013.2269186>
- [253] T.-H. Tsai, W.-H. Cheng, and Y.-H. Hsieh. 2011. Dynamic social network for narrative video analysis. In *19th ACM international conference on Multimedia*. 663–666. <https://doi.org/10.1145/2072298.2072413>
- [254] H. Vala, S. Dimitrov, D. Jurgens, A. Piper, and D. Ruths. 2016. Annotating Characters in Literary Corpora: A Scheme, the Charles Tool, and an Annotated Novel. In *10th International Conference on Language Resources and Evaluation*. 184–189. http://www.lrec-conf.org/proceedings/lrec2016/pdf/1130_Paper.pdf
- [255] H. Vala, D. Jurgens, A. Piper, and D. Ruths. 2015. Mr. Bennet, his coachman, and the Archbishop walk into a bar but only one of them gets recognized: On The Difficulty of Detecting Characters in Literary Texts. In *Conference on Empirical Methods in Natural Language Processing*. http://piperlab.mcgill.ca/pdfs/Vala_EMNLP_2015.pdf
- [256] J. Valls-Vargas, S. Ontañón, and J. Zhu. 2014. Toward automatic character identification in unannotated narrative text. In *7th Intelligent Narrative Technologies Workshop*. 188–194. <https://pdfs.semanticscholar.org/b755/1e222321cca2fdb1992991c434ef0aa1a2fd.pdf>
- [257] K. van Dalen-Oskam, J. de Does, M. Marx, I. Sijaranamual, K. Depuydt, B. Verheij, and V. Geirnaert. 2014. Named entity recognition and resolution for literary studies. *Computational Linguistics in the Netherlands Journal* 4 (2014),

- 121–136. <https://clinjournal.org/node/62>
- [258] M. van de Camp and A. van den Bosch. 2011. A link to the past: constructing historical social networks. In *2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*. 61–69. <https://dl.acm.org/citation.cfm?id=2107661>
- [259] T. Venturini, L. Bounegru, M. Jacomy, and J. Gray. 2016. How to Tell Stories with Networks: Exploring the Narrative Affordances of Graphs with the Iliad. In *Datafied Society: Social Research in the Age of Big Data*. Amsterdam University Press. http://www.tommasoventurini.it/wp/wp-content/uploads/2016/01/How_to_Tell_Stories_with_Networks_PreprintVersion.pdf
- [260] A. Vinciarelli, F. Fernandez, and S. Favre. 2007. Semantic Segmentation of Radio Programs using Social Network Analysis and Duration Distribution Modeling. In *IEEE International Conference on Multimedia and Expo*. 779–782. <https://doi.org/10.1109/ICME.2007.4284766>
- [261] A. V. Voloshinov and I. V. Gozhanskaya. 2008. Russian vs. English Drama in the Context of Network Theory. In *XXth International Association of Empirical Aesthetics Biennial Congress*. 19–22. http://ets.ifmo.ru/tomasov/konferenc/AutoPlay/Docs/Volume%204/8_06.pdf
- [262] M. Vrigkas, C. Nikou, and I. A. Kakadiaris. 2015. A Review of Human Activity Recognition Methods. *Frontiers in Robotics and AI* 2 (2015). <https://doi.org/10.3389/frobt.2015.00028>
- [263] D. J. Watts and S. H. Strogatz. 1998. Collective dynamics of ‘small-world’ networks. *Nature* 393, 6684 (1998), 440–442. <https://doi.org/10.1038/30918>
- [264] M. C. Waumans, T. Nicodème, and H. Bersini. 2015. Topology Analysis of Social Networks Extracted from Literature. *PLoS ONE* 10, 6 (2015), e0126470. <https://doi.org/10.1371/journal.pone.0126470>
- [265] D. Weinland, R. Ronfard, and E. Boyer. 2011. A survey of vision-based methods for action representation, segmentation and recognition. *Computer Vision and Image Understanding* 115, 2 (2011), 224–241. <https://doi.org/10.1016/j.cviu.2010.10.002>
- [266] G. Weissman. 2011. Grey’s Anatomy Network of Sexual Relations. <https://gweissman.github.io/babelgraph/blog/2017/06/15/greys-anatomy-network-of-sexual-relations.html>. (03 2011).
- [267] C.-Y. Weng, W.-T. Chu, and J.-L. Wu. 2007. Movie analysis based on roles’ social network. In *IEEE International Conference on Multimedia and Expo*. 1403–1406. <https://doi.org/10.1109/ICME.2007.4284922>
- [268] C.-Y. Weng, W.-T. Chu, and J.-L. Wu. 2007. RoleNet: Treat a movie as a small society. In *International Workshop on multimedia information retrieval*. 51–60. <https://doi.org/10.1145/1290082.1290092>
- [269] C.-Y. Weng, W.-T. Chu, and J.-L. Wu. 2009. RoleNet: Movie Analysis from the Perspective of Social Networks. *IEEE Transactions on Multimedia* 11, 2 (2009), 256–271. <https://doi.org/10.1109/TMM.2008.2009684>
- [270] P. Wills and F. G. Meyer. 2019. Metrics for Graph Comparison: A Practitioner’s Guide. *arXiv stat.AP* (2019), 1904.07414. <https://arxiv.org/abs/1904.07414>
- [271] A. Woloch. 2003. *The One vs. the Many: Minor Characters and the Space of the Protagonist in the Novel*. Princeton University Press. <http://press.princeton.edu/titles/7622.html>
- [272] A. Xanthos, I. Pante, Y. Rochat, and M. Grandjean. 2016. Visualising the dynamics of character networks. In *Digital Humanities*. 417–419. https://github.com/maladesimaginaires/intnetviz/raw/master/DH2016_xanthos_et_al.pdf
- [273] M.-C. Yeh, M.-C. Tseng, and W.-P. Wu. 2012. Automatic social network construction from movies using film-editing cues. In *IEEE International Conference on Multimedia and Expo Workshops*. 242–247. <https://doi.org/10.1109/ICMEW.2012.48>
- [274] M.-C. Yeh and W.-P. Wu. 2014. Clustering Faces in Movies Using an Automatically Constructed Social Network. *IEEE MultiMedia* 21, 2 (2014), 22–31. <https://doi.org/10.1109/MMUL.2014.24>
- [275] C. Y. Yeung and J. Lee. 2017. Identifying Speakers and Listeners of Quoted Speech in Literary Works. In *8th International Joint Conference on Natural Language Processing*. 325–329. <http://www.aclweb.org/anthology/I17-2055>
- [276] M. Yeung, B.-L. Yeo, and B. Liu. 1996. Extracting story units from long programs for video browsing and navigation. In *3rd IEEE International Conference Multimedia Computing & Systems*. 296–305. <https://doi.org/10.1109/MMCS.1996.534991>
- [277] J. Yose, R. Kenna, M. MacCarron, and P. MacCarron. 2018. Network Analysis of the Viking Age in Ireland as portrayed in Cogadh Gaedhel re Gallaibh. *Royal Society Open Science* 5 (2018), 171024. <https://doi.org/10.1098/rsos.171024>
- [278] J. Yose, R. Kenna, P. MacCarron, T. Platini, and J. Tonra. 2016. A Networks-Science Investigation into the Epic Poems of Ossian. *Advances in Complex Systems* 19 (2016), 1650008. <https://doi.org/10.1142/S0219525916500089>
- [279] K. Yuan, R. Ji, H. Yao, X. Sun, P. Xu, and X. Liu. 2009. VisualCor system: search actor correlations in TV series. In *1st International Conference on Internet Multimedia Computing and Service*. 213–218. <https://doi.org/10.1145/1734605.1734655>
- [280] K. Yuan, H. Yao, R. Ji, and X. Sun. 2010. Mining actor correlations with hierarchical concurrence parsing. In *IEEE International Conference on Acoustics, Speech and Signal Processing*. 798–801. <https://doi.org/10.1109/icassp.2010.5494953>

- [281] J. Y. Zhang, A. W. Black, and R. Sproat. 2003. Identifying Speakers in Children’s Stories for Speech Synthesis. In *Eurospeech*. 2041–2044. <https://www.cs.cmu.edu/~awb/papers/eurospeech2003/esper.pdf>
- [282] Y.-F. Zhang, C.-S. Xu, H.Q. Lu, and Y.-M. Huang. 2009. Character Identification in Feature-Length Films Using Global Face-Name Matching. *IEEE Transactions on Multimedia* 11, 7 (2009), 1276–1288. <https://doi.org/10.1109/TMM.2009.2030629>