



Pamphlet **2**
→ May 1, 2011

Literary **Lab**

Network Theory,
Plot Analysis

Franco Moretti

Network Theory, Plot Analysis

In the last few years, literary studies have experienced what we could call the rise of quantitative evidence.¹ This had happened before of course, without producing lasting effects, but this time it's probably going to be different, because this time we have digital databases, and automated data retrieval. As Michel's and Lieberman's recent article on "Culturomics" made clear, the width of the corpus and the speed of the search have increased beyond all expectations: today, we can replicate in a few minutes investigations that took a giant like Leo Spitzer months and years of work.² When it comes to phenomena of language and style, we can do things that previous generations could only dream of.

When it comes to language and style. But if you work on novels or plays, style is only part of the picture. What about plot – how can that be quantified? This paper is the beginning of an answer, and the beginning of the beginning is network theory. This is a theory that studies connections within large groups of objects: the objects can be just about anything – banks, neurons, film actors, research papers, friends... – and are usually called nodes or vertices; their connections are usually called edges; and the analysis of how vertices are linked by edges has revealed many unexpected features of large systems, the most famous one being the so-called "small-world" property, or "six degrees of separation": the uncanny rapidity with which one can reach any vertex in the network from any other vertex. The theory proper requires a level of mathematical intelligence which I unfortunately lack; and it typically uses vast quantities of data which will also be missing from my paper. But this is only the first in a series of studies we're doing at the Stanford Literary Lab; and then, even at this early stage, a few things emerge.

¹ The text of this pamphlet has appeared – in a somewhat abbreviated form – in *New Left Review* 68, March-April 2011. Our thanks to NLR for permission to post the pamphlet on the Literary Lab's website.

² Jean-Baptiste Michel, Erez Lieberman Aiden, et al., "Quantitative Analysis of Culture Using Millions of Digitized Books", *Science*, December 2010.

1. Character-network

A network is made of vertices and edges; a plot, of characters and actions: characters will be the vertices of the network, interactions the edges, and here is what the *Hamlet* network looks like: **Figure 1**. There are some questionable decisions here, mostly about *The Murder of Gonzago*, but, basically, two characters are linked if some words have passed between them: an interaction, is a speech act. This is not the only way to do things, the authors of a previous paper on Shakespeare had linked characters if they had speaking parts during the same scene, even if they did not address each other: so for instance, for them the Queen and Osric are linked (because they both have speaking parts, and are on stage together in the last scene of the play), whereas here they are not, because they don't speak to each other.³ My network uses explicit connections, theirs adds implicit ones, and is obviously denser, because it has all of my edges plus some; both are plausible, and both have at least two flaws. First, the edges are not "weighted": when Claudius tells Horatio in the graveyard scene, "I pray thee, good Horatio, wait upon him", these eight words have in this Figure exactly the same value as the four *thousand* words exchanged between Hamlet and Horatio. This can't be right. And then, the edges have no "direction": when Horatio addresses the Ghost in the opening scene, his words place an edge between them, but of course that the Ghost wouldn't reply and would only speak to Hamlet is important, and should be made visible.⁴ But, I just couldn't find a non-clumsy way to visualize weight and direction; and as a consequence, the networks in this study were all made by hand, with the very simple aim of maximizing visibility by minimizing overlap. This is not a long term solution, of course, but these are small networks, where intuition can still play a role; they're like the childhood of network theory for literature; a brief happiness, before the stern adulthood of statistics.

Anyway. Four hours of action, that become this. Time turned into space: a character-system arising out of many character-spaces, to use Alex Woloch's concepts in *The One vs. the Many*. Hamlet's space, (**Figure 2**), defined as the set of characters he interacts with directly; Hamlet and Claudius (**Figure 3**): see how much of the network they capture, between the two of them. Ophelia and Gertrude (**Figure 4**): the much smaller space of the two women in the play. And so on. But before analyzing spaces in detail, why use networks to think about plot to begin with? What do we gain, by turning time into space? First of all, this: when we watch a play, we are always in the present: what is on stage, is; and then it disappears. Here, nothing ever disappears. What is done, cannot be undone. Once the

³ "The network structure calculations were obtained by treating each speaking character as a vertex, and deeming two characters to be linked if there was at least one time slice of the play in which both were present (that is, if two characters spoke to each other or were in each other's presence, then they have a link)": James Stiller, Daniel Nettle, Robin I.M. Dunbar, "The small world of Shakespeare's plays", *Human Nature*, vol. 14, no. 4, 2003, p. 399. Another application of network theory to narrative ("Marvel Universe looks almost like a real social network", by R. Alberich, J. Miro-Julia, F. Rossello, <http://arxiv.org/abs/cond-mat/0202174v1>) uses a similar premise, by stating that "two characters are linked when they jointly appear in a significant way in the same comic book"; since, however, we are never told what exactly constitutes a "significant" interaction, as opposed to an in-significant one, the basis for quantification remains fundamentally opaque.

⁴ The reason weight and direction are particularly important in literary networks is that, whereas the systems studied by network theory have easily thousands or millions of vertices, whose relevance can be directly expressed in the number of connections, plots have usually no more than a few dozens characters; as a consequence, the mere existence of a connection is seldom sufficient to establish a hierarchy, and must be integrated with other measurements.

Ghost shows up at Elsinore things change forever, whether he is on scene or not, because he is never not there in the network. The past becomes past, yes, but it never disappears from our perception of the plot.

Making the past just as visible as the present: that is one major change introduced by the use of networks. Then, they make visible specific “regions” within the plot as a whole: sub-systems, that share some significant property. Take the characters who are connected to both Claudius and Hamlet in **Figure 5**: except for Osric and Horatio, whose link to Claudius is however extremely tenuous, they are all killed. Killed by whom, is not always easy to say: Polonius is killed by Hamlet, for instance – but Hamlet has no idea that it is Polonius he is stabbing behind the arras; Gertrude is killed by Claudius – but with poison prepared for Hamlet, not for her; Hamlet is killed by Laertes, with Claudius’s help, while Laertes himself, like Rosencrantz and Guildenstern before him, are all killed by Hamlet, but with Claudius’s weapons. Individual agency is muddled; what is truly deadly, is the characters’ position in the network, chained to the warring poles of king and prince. Outside of that red region, no one dies in *Hamlet*. The tragedy, is all there.

2. Models, experiments

Third consequence of this approach: once you make a network of a play, you stop working on the play proper, and work on a *model* instead: you reduce the text to characters and interactions, abstract them from everything else, and this process of reduction and abstraction makes the model obviously much less than the original object – just think of this: I am discussing *Hamlet*, and saying nothing about Shakespeare’s words – but also, in another sense, much *more* than it, because a model allows you to see the underlying structures of a complex object. It’s like an X-ray: suddenly, you see the region of death of **Figure 5**, which is otherwise hidden by the very richness of the play. Or take the protagonist. When discussing this figure, literary theory usually turns to concepts of “consciousness” and “interiority” – even Woloch’s structural study takes this path. When a group of researchers applied network theory to the Marvel series, however, their view of the protagonist made no reference to interiority; the protagonist was simply “the character that minimized the sum of the distances to all other vertices”;⁵ in other words, the *center* of the network. In their case, it was a character called Captain America; in ours, it’s Hamlet. One degree of separation from 16 of the characters (**Figure 6**); two degrees from the others (**Figure 7**); average distance from all vertices in the network, 1.45. And if we visualize these results in the form of a histogram, (**Figure 8**), we find the power-law distribution that is characteristic of all networks: very few characters with many edges on the left, and very many characters with just one or two edges on the right; and the result is the same if we add all the characters from *Macbeth*, *Lear*, and *Othello* (**Figure 9**). Power-law is the opposite of a Gaussian curve: there is no central tendency in the distribution, no “average”; that is to say, there is no “typical” vertex in the network, and *no typical character in the plays*. So, speaking of Shakespeare’s characters “in general” is wrong, at least in the tragedies, because these characters-in-general don’t exist: all there is, is this curve leading from one extreme to the other without any clear solution of continuity. And the same applies to the bina-

⁵ <http://arxiv.org/abs/cond-mat/0202174v1>

ries with which we usually think about character: protagonist versus minor characters, or “round” versus “flat”: nothing in the distribution supports these dichotomies; what it asks for, rather, is a radical reconceptualization of characters and of their hierarchy.

What is done is never undone; the plot as a system of regions; the hierarchy of centrality that exists among characters; finally – and it’s the most important thing of all, but also the most difficult – one can *intervene* on a model; make experiments. Take the protagonist again. For literary critics, this Figure is important because it’s a very meaningful part of the text; there is always a lot to be said about it; we would never think of discussing *Hamlet* – without Hamlet. But this is exactly what network theory tempts us to do: take the *Hamlet*-network (**Figure 10**), and remove Hamlet, to see what happens: **Figure 11**. And what happens is that the network almost splits in half: between the court on the right, and the region that includes the Ghost and Fortinbras on the left all that remains are the three edges linking Horatio to Claudius, Gertrude, and Osric: a few dozen words. If we used the first Quarto, the breakdown would be even more dramatic: **Figures 12-3**.

Why is the protagonist significant here? Not for what is “in” it; not for its essence, but for its function in the stability of the network. And stability has clearly much to do with centrality, but is not identical to it. Take the second most central character of the play: Claudius. In quantitative terms, Claudius is almost as central as Hamlet (average distance of 1.62, versus 1.45); but in structural terms not so, when we remove him from the network (**Figures 14-5**) what happens is that a handful of peripheral characters are affected, but the network as a whole not much. Even if we remove, first Hamlet, and then Claudius (**Figures 16-8**), his subtraction doesn’t do much. But if we remove, first Hamlet, and then *Horatio* (**Figures 19-21**), then the fragmentation is so radical that the Ghost and Fortinbras – which is to say, the beginning and the ending of the play – are completely severed from each other and from the rest of the plot. *Hamlet* no longer exists. And yet, Horatio is slightly less central than Claudius in quantitative terms (1.69 versus 1.62). Why is he so much more important in structural terms?

3. Centrality, conflict, clustering

Let me take a brief step back, and add something on Hamlet’s centrality first. Shakespeare’s major tragedies are reflections on the nature of sovereignty, in which an initial Figure of legitimacy is ousted by an usurper, who is in his turn defeated by a second Figure of legitimacy. But there are differences. In *Macbeth* and *Lear* legitimate rulers have very solid connections to the rest of the network: Duncan and Malcolm (**in red and purple: Figure 22**) have a powerful antagonist in Macbeth (**green: Figure 23**), but the two fields are basically balanced; and this is even truer for *Lear*, with its scattering of sovereign power (**Figures 24-5**). In *Hamlet*, no: between old Hamlet and Fortinbras on one side, and Claudius on the other (**Figures 26-27**) there is a total disproportion: the usual balance of power is not there,⁶ and Hamlet finds himself caught between the space of the Court, and that of the anti-Court: the soldiers who still remember the old king, the ghost, the Norwe-

⁶ Why the balance is not there – why choose a ghost and a Norwegian as Figures of legitimacy – is a different question, on which network theory has probably nothing to say. *That* it is not there, is one of those things that it makes visible.

gian pretender, the carnivalesque of the Gravedigger. It's a duality that emerges in all the great Court scenes, from that which sets up the pattern in act I (**Figure 28**), to the arrival of the players (**Figure 29**), the play within the play (**Figure 30**), and the two final scenes of the tragedy (**Figures 31-2**). Always two hubs in the network: Claudius inside the Court, and Hamlet (half-)outside it.

Claudius inside the Court ... This is the densest part of the network: the hexagon formed by Hamlet, Claudius, Gertrude, Polonius, Ophelia and Laertes (**Figure 33**), where everybody is connected to everybody else, and clustering reaches 100%. Clustering is a technical concept of network theory, which Mark Newman explains thus: "If vertex A is connected to vertex B and vertex B to vertex C, then there is a heightened probability that vertex A will also be connected to vertex C. In the language of social networks, the friend of your friend is likely also to be your friend."⁷ This is what clustering means: A and C connect, the triangle closes, and when that happens the resilience of that part of the network increases. *And this is why is removing Claudius has such little effect on the network*: he belongs to a region which is already very interconnected, and that remains just as solid with or without him.⁸

Horatio is the opposite: he inhabits a part of the network where clustering is so low (**Figure 34**) that, without him, it disintegrates. In this, he is a good gateway to the region that is the exact antithesis of the 100% clustering of the Court: the periphery of *beyond*, where we find the least connected of its characters (**Figure 35**): those with just one link to the network; at times, just one sentence. Very little. But as a group, these peripheral characters do something unique: they point to the world *beyond Elsinore*: the gentleman, sailor and ambassadors who speak to Horatio, and one of the messengers to Claudius, are links to the "English" subplot; Cornelius and Voltmand, to "Norway"; Reynaldo, to Laertes's "France"; the Priest and Grave-digger, to the world of the dead. These centrifugal threads – "tendrils", as they're sometimes called – contribute to the uncanny feeling that Elsinore is just the tip of the tragic iceberg: geography as the hidden dimension of fate, like genealogy in Greek tragedy. Genealogy, vertical, rooted in myth; geography, horizontal, in something like the nascent European state system.

7 Mark Newman, "The Structure and Function of Complex Networks", *SIAM Review*, vol. 45. no. 2, 2003, p. 183.

8 Hamlet also belongs to the hexagon, of course; but although he shares those five edges with Claudius (plus that to Horatio, and to those other Court creatures, Rosencrantz, Guildenstern, and Osric), their remaining edges are quite different: in Claudius's case, they link him to minor characters who are emanations of the Court, and hence add nothing to his role in the structure; in Hamlet's case, they lead into other regions of the play, increasing his structural significance. Moreover, whereas Hamlet's exchanges with the five Court characters amount only to 28% of the words he speaks in the play, in Claudius's case – though he barely speaks to Ophelia, and not much to Polonius, either – the figure rises to 48% (or 60%, if we include his speeches to the Court as a whole): in other words, most of Claudius's verbal energy is spent within this very small circle. This is one case where "weighting" the edges would significantly modify the initial X-ray of *Hamlet*.

4. Horatio

I may be exaggerating here, projecting onto the periphery of this diagram Napoleon's words at Erfurt on politics as the fate of the moderns. But Horatio's space – ambassadors, messengers, sentinels, talk of foreign wars, and of course the transfer of sovereignty at the end – all this announces what will be soon called, not Court, but State. The Court, the space of 100% clustering, where one is always seeing and being seen, as in Elias, is really two families: Ophelia, Laertes, and Polonius; Claudius, Gertrude, and Hamlet. Horatio's world is more abstract: he exchanges just a couple of sentences with Claudius and Gertrude, and none at all with Polonius, Ophelia and Laertes. Here, incidentally, you see the difference between my network and that other Shakespeare study: for them Horatio is linked to Polonius Laertes and Ophelia, because they are on scene together, which seems to me to miss the point of his character: his being a “weak tie”, unlike those hyperconnected families-at-Court. Weak, that is to say: less intense, but with a wider radius; and more impersonal, almost bureaucratic, like the ties described by Graham Sack in his study of *Bleak House*.⁹

I may be making too much of this; or, Horatio may really be a fantastic half-intuition on Shakespeare's part; and I say “half”, because there is something enigmatically undeveloped about him. Think of Posa, in Schiller. *Don Carlos* is to a large extent a remake of *Hamlet*, and Posa is certainly a remake of Horatio: another lonely friend of another sad prince in another oedipal play. But Posa has a reason for being so central: he is that new Figure, so important for modern drama: the ideologue. There is something he wants to *do*. Horatio? Kent is near Lear out of loyalty; Macduff, near Malcolm to avenge his family. Horatio?

Horatio has a function in the play, but not a motivation. No aim, no emotions – no *language*, really, worthy of *Hamlet*. I can think of no other character that is so central to a Shakespeare play, and so flat in its style. Flat, just like the style of the State (or at least, of its bureaucracy). Flat, like the typical utterances we encounter at the periphery of *Hamlet*: orders, and news: “And we here dispatch/ You, good Cornelius, and you, Voltmand” (I.2.33-4); “Sea-faring men, sir. They say they have letters for you” (IV.6.2-3). Orders and news must avoid ambiguity, and so, around them, the play's “Figureality rate” (to use a concept of Francesco Orlando's) drops; language becomes simple. Conversely, as we move towards the center of the network Figureality rises, all the way to Hamlet's puns in response to Claudius, and to the soliloquies that occupy, so to speak, the center of the center. You see the possibility here, which we are pursuing in a follow-up study: different uses of language emerging in different network regions. Style, integrated within plot as a *function* of plot. It would be a breakthrough, and not just for literary analysis – which has never been able to create a unified theory of plot and style – but for the analysis of culture more broadly: because plot and style could provide a small-scale model to study two general properties of human societies: plot, to understand how the simple exchange between two individuals evolves into complex patterns made of thousands of interactions; and style, to study how human beings make sense of their actions. A model for the relationship between what we do, and how we think about it: this is what a plot-style continuum could provide. But we are definitely not there yet.

⁹ Alexander Graham Sack, “*Bleak House* and Weak Social Networks”, unpublished thesis, Columbia University, 2006. The concept of “weak tie” was first formulated by Mark Granovetter in “The Strength of Weak Ties”, *American Journal of Sociology*, May 1973.

5. Symmetry

Networks are made of vertices and edges; plot networks, of characters and verbal exchanges. In plays this works well, because words are deeds, deeds are almost always words, and so, basically, a network of speech acts is a network of actions. In novels, no, because much of what characters do and say is not uttered, but narrated, and direct discourse covers only a part of the plot – at times, a very small part. This makes the transformation of plots into networks a lot less accurate, but the idea is too tempting to just let it go, and so I will show a few networks of verbal exchanges from *The Story of the Stone* and *Our Mutual Friend* just the same. A couple of years ago I conjectured that the number of characters could be a major source of morphological differences between Chinese and Western novels, and networks seem to be a good way to test the idea.

Unlike *Hamlet* however, I won't present networks for the entire text, but only chapter-networks; I could perhaps manage *Our Mutual Friend* (even though, by Western standards, it has a lot of characters), but certainly not the hundreds and hundreds of characters of *The Story of the Stone*, where each chapter has between 5 and 28 different speaking characters, with a median of 14. *Our Mutual Friend* is less crowded: between 3 and 14 speaking characters per chapter, with a median of 6. And here is one of them: chapter one of book two of the novel, (**Figure 36**), which introduces Jenny Wren and Headstone; chapter two, a variation on this, (**Figure 37**), with Lizzie's other suitor, Wrayburn, and Jenny's father; four, with the revenge of the Lammlers over Podsnap via his daughter (**Figure 38**). And so on.

Now, in Western novelistic poetics, aside from a few neo-classical moments, symmetry has never been an important category. But you look at the networks from *Our Mutual Friend* – (**Figures 39-45**) – and it's stunning how regular they are. Probably, there are two reasons for this. The first is that Dickens's building blocks are usually binary pairs: husband and wife, parent and child, brother and sister, suitor and beloved, friend and friend, employer and employee, rival and rival ... And, second, these binaries can project their dualism onto the chapter as a whole because there is very little "noise" around them – very few other characters to disrupt the symmetry. Or in other words: with few characters, symmetry seems to emerge by itself, even in the absence of an aesthetics of symmetry.

An aesthetics of symmetry is on the other hand very present in Chinese literary culture, where readers of novels expect, writes Andrew Plaks, that "the overall sequence of chapters (...) will add up to a round and symmetrical number, typically 100 or 120. The pronounced sense of symmetry (...) provides the ground for a variety of exercises in structural patterning. Most noticeable among these is the practice of contriving to divide an overall narrative sequence precisely at its arithmetic midpoint, yielding two great hemispheric structural movements."¹⁰ Hemispheric movements ... Think of the rhymed couplets that serve as chapter epigraphs in classical Chinese novels: "Zhou Rui's wife delivers palace flowers and finds Jia Lian pursuing night sports by day / Jia Bao-yu visits the Ning-guo mansion and has an agreeable colloquy with Qin-shi's brother". A does this and meets B; C does that and meets D. As if the two halves of the chapter mirrored each other perfectly: "A very earnest young woman offers counsel by night / And a very endearing one is found to be a source of fragrance by day". "Parallel prose", as Chinese aesthetics calls it. So you take *The Story of the Stone*, use black edges for the first half of the chapter, red edges for the second half, and ... **Figures 46-52**.

10 Andrew Plaks, "The Novel in Premodern China", in Franco Moretti, ed., *The Novel*, Princeton UP 2006, vol. I, p. 189.

Chinese novels should have *more* symmetry than European ones. But no. And the number of characters is probably again the reason: if with few characters symmetry emerges almost by itself, with *many* characters it becomes implausible. It's one of those cases where size is not just size: it's form. But what does this form mean? Dickens's symmetry is clear: it indicates that, below the surface of social interactions, there is always a melodramatic substratum of love or hatred ready to erupt. A-symmetry?

6. *Guanxi*

First half of the seventh chapter of *The Story of the Stone* (Figure 53). Zhou Rui's wife, who is a member of the staff of the Rong mansion, must report to Lady Wang on the visit of a distant relative; she doesn't find her in her apartment, asks about her, is sent to other parts of the compound, is given some errands, inquires about some new faces and about people she hasn't seen in a while, is asked to intercede for her son-in-law ... and so she ends up meeting a dozen characters – or more exactly, *speaking* to a dozen characters, she meets about twice as many, while another twenty or so are mentioned in the various conversations.

Nothing major happens here: people talk, walk around, play *go*, gossip... No interaction is crucial in itself. But taken together, they perform an essential reconnaissance function: they make sure that the nodes in this region are still communicating: because, with hundreds of characters, the disaggregation of the network is always a possibility. We are close to one of the most distinctive keywords of Chinese culture: *guanxi*: something like “connections”, translate Gold, Guthrie and Wank; part of “a specifically Chinese idiom of social networks (...) linked to other building blocks of sociality such as *ganqing* (sentiment), *renqing* (human feelings), *mianzi* (face), and *bao* (reciprocity)”: a world which is “neither individual- nor society-based, but *relation*-based”.¹¹ And these relations are not a given, they are an artifact; “manufacturing obligation”, “chain of transactions”, “indebtedness”, “consciously producing” connections – this is the lexicon of *guanxi*.¹²

A chain of transactions that generate indebtedness: in chapter 24 of the novel (Figure 54), Jia Yun, who is a poor relative of the Rong-guo house, is looking for work; he asks Jia Lian, receives only vague promises, so he turns to his uncle Bu Shi-ren, who owns a store, hoping to get some perfumes on credit to use as presents. Bu Shi-ren says no, Jia Yun walks away and bumps into a drunk, who turns out to be his neighbor Ni Er, a racketeer; Ni Er finally lends him the money, and Jia Yun buys a present for Xi-feng, who is in charge of the finances of the clan. This is how *guanxi* works – and this is what creates the asymmetry: a character rallies all its resources in order to “manufacture obligation”, unbalancing

11 Thomas Gold, Doug Guthrie and David Wank, “An Introduction to the Study of *Guanxi*”, in *Social Connections in China. Institutions, Culture, and the Changing Nature of Guanxi*, Cambridge UP 2002, pp. 3, 4, 10.

12 See Gold, Guthrie and Wank, cit., p. 6, Mayfair Mei-hui Yang, *Gifts, Favours and Banquets. The art of social relationships in China*, Cornell UP 1994, pp. 6, 44, and 125, and Andrew Kipnis, “Practices of *guanxi* production and practices of *ganqing* avoidance”, in *Social Connections in China*, cit..

a whole cluster of interactions in the same direction. Ideally, in the long run *guanxi* will produce reciprocity, and hence symmetry: but at the scale of *the chapter*, asymmetry is exactly what we should expect. And, needless to say, a story which is unbalanced at the local scale, and balanced at a higher one – this is interesting. Even more so, if in Dickens we were to find the opposite configuration: symmetry in the chapters – and asymmetry in the plot as a whole. We'll see.

7. Fruitful doing

In the last two Figures, I have focused on how individual behaviour contributes to the shape of the network; now I'll turn the matter around, to see how the overall network of *The Story of the Stone* shapes individual characters in a specific way. Bao-yu, in chapter 8, is a good instance of this (Figure 55): as the chapter unfolds, he takes part in three distinct episodes: he has an important encounter with his pre-destined bride Bao-chai, which is catalyzed by her maid Oriole; then he gets drunk amidst the banter of the characters around him, despite Nannie Li's vigilance; finally, he throws a tantrum with his maids, until Aroma threatens a general desertion. Three episodes; all mediated by different characters; each of them bringing out a distinct side of Bao-yu (naive lover, sensuous youth, petty domestic tyrant) due to his interaction with a different cluster of characters. And the same happens in every chapter of the novel: its huge pack of characters is re-shuffled, the new "hand" forms new character-clusters, which generate new features in the Figures we already knew. Novelty, as the result of recombination: in the first twenty chapters of the novel, Bao-yu speaks to 54 characters, and not once does the same group reform around him.

Now, Bao-yu is arguably the protagonist of *The Story of the Stone*: the male child born under very special auspices, and expected to do great things for his family. But what a strange life, for a protagonist: constantly summoned by this and that relative, kept under supervision, asked to perform all sorts of duties – even the many delightful opportunities he is offered come usually with constraints attached. The protagonist, yes, but not free. The protagonist, and therefore not free: because he has a duty *towards the structure*: towards the relation-based society he is part of. "The One *Bildungsroman* the Many": Elizabeth Bennet, not off to Pemberley on her own, but kept at home, to shape the life of her sisters.

A different role for the protagonist, resulting from a different set of narrative relations: what networks make visible are the opposite foundations of novel-writing East and West. One day, after we add to these skeletons the layers of direction, weight, and semantics, those richer images will perhaps make us see different genres – tragedies and comedies; picaresque, gothic, *Bildungsroman* ... – as different shapes; ideally, they may even make visible the micro-patterns out of which these larger network shapes emerge. But for this to happen, an enormous amount of empirical data must be first put together. Will we, as a discipline, be capable of sharing raw materials, evidence – *facts* – with each other? It remains to be seen. For science, Stephen Jay Gould once wrote, fruitful doing matters more than clever thinking. For us, not yet.

8. Epilogue. Theory, visualization, concepts

The idea behind this study, clearly stated in its opening page, was, very simply, that network theory could offer a way to quantify plot, thus providing an essential piece that was still missing from computational analyses of literature. Once I started working in earnest, though, I soon realized that the machine-gathering of the data, essential to large-scale quantification, was not yet a realistic possibility. (Others, elsewhere, were already at work on this problem; but I wasn't aware of it). So, from its very first section, the essay drifted from quantification to the qualitative analysis of plot: the advantage of thinking in terms of space rather than time; its segmentation into regions, instead of episodes; the new, non-anthropomorphic idea of the protagonist; or, even, the "undoing" of narrative structures occasioned by the removal of specific vertices in the network.

Looking back at the work done, I wouldn't call this change of direction a mistake: after all, network theory *does* help us redefine some key aspects of the theory of plot, which is an important aspect of literary study. This is not the theory's original aim, of course, but then again, a change of purpose – a "refunctionalization", as the Russian Formalists called it – is often what happens to a system of thought traveling from one discipline to another. Within this first change of direction, however, a second, more radical one occurred. Yes, my analysis took network theory as its starting point, and in a couple of cases (the pages on clustering and *guanxi*) it also engaged its categories; but, by and large, my objects and concepts were quite distant from them. The "region of death"; the "spaces" of legitimacy and usurpation; "symmetry"; none of these had much to do with network theory. As I have often been asked when presenting the paper in public: Did I really need it, to speak about Horatio and the State?

No, I did not need network *theory*; but I probably needed *networks*. I had been thinking about Horatio for some time – but I had never "seen" his position within *Hamlet's* field of forces until I looked at the network of the play. "Seen" is the keyword here. What I took from network theory were less concepts than *visualization*: the possibility of extracting characters and interactions from a dramatic structure, and turning them into a set of signs that I could see at a glance, in a two-dimensional space. Basically, I used (or mis-used) the theory in the same way I had used cartography in the *Atlas of the European Novel*, and charts in *Graphs, Maps, Trees*: as a way of arranging literary data that presupposed a principle of order – but not a full conceptual architecture.

My fondness for the "intermediate" epistemological status of visualization reflects, I suspect, my general position towards scientific models: always impressed by their explanatory power, I am woefully unprepared to use them properly, especially once math enters the picture (that is to say, very soon). Whence this study's unplanned, yet crucial, decisions to replace quantification with qualitative analysis, and then basing the latter less on concepts than on my own intuitions vis-à-vis visual evidence. It's the note sounded by the final paragraph of the essay, with its hope that, one day, "richer images will perhaps make us see different genres (...) as different shapes; ideally, even make visible the micro-patterns out of which these larger network shapes emerge." Images, see, visible, patterns, shapes ... The passage is saturated by the lexicon of visualization, projected into the distant future.

And so, a few months ago, the follow-up study began indeed by testing the “visibility” of the (new) network patterns based on machine-gathered data for all of Shakespeare, and elaborated by networking programs. The details of the new experiment – beginning with the algorithm designed by Matt Jockers to identify the addressees of dramatic utterances – will be described in Rhiannon Lewis’s forthcoming study. Here, let me end this brief coda with a vignette of the Lab meeting which marked a turning point in our work. Lewis, Jockers, Heuser and me were discussing the new images, that had been tirelessly tweaked by Jockers and Heuser to maximize our chances of “seeing” the complex dramatic structures we wanted to understand. We had all worked hard at this; had discarded many disappointing solutions; and were now concentrating on the most synthetic, most “improved” versions. For the sake of continuity with the present pamphlet, let me just show the two new images of the *Hamlet* network: **Figures 56-7**.

Now, there is no question that these figures contained much more information than Figure 1 of the pamphlet: they showed, not just who had talked to whom, but also whether the exchange had been mutual, and how extensive it had been (measured in the number of words). All this was new. Was it also *visible*? Clearly the answer was no. The images were much more confusing, more impervious to intuition. And, remember, they were the best of the group.

It is never easy, realizing that one has reached a dead end, pure and simple. But this is what it was. Using networks to gain intuitive knowledge of plot structures had played an important role – but we had now reached the limits of its usefulness. Better turn away from images for a while, and let intuition give way to concepts (network size, density, clustering, betweenness ...), and to statistical analysis. And so, the dead end of this pamphlet, was the beginning of Rhiannon Lewis’s work.

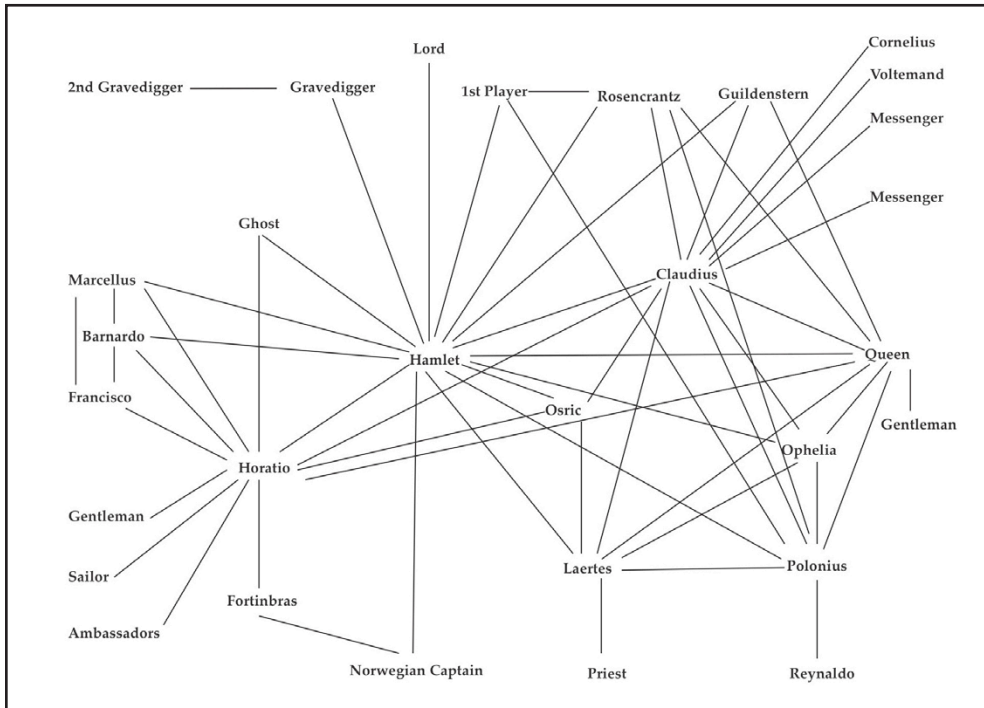


Figure 1

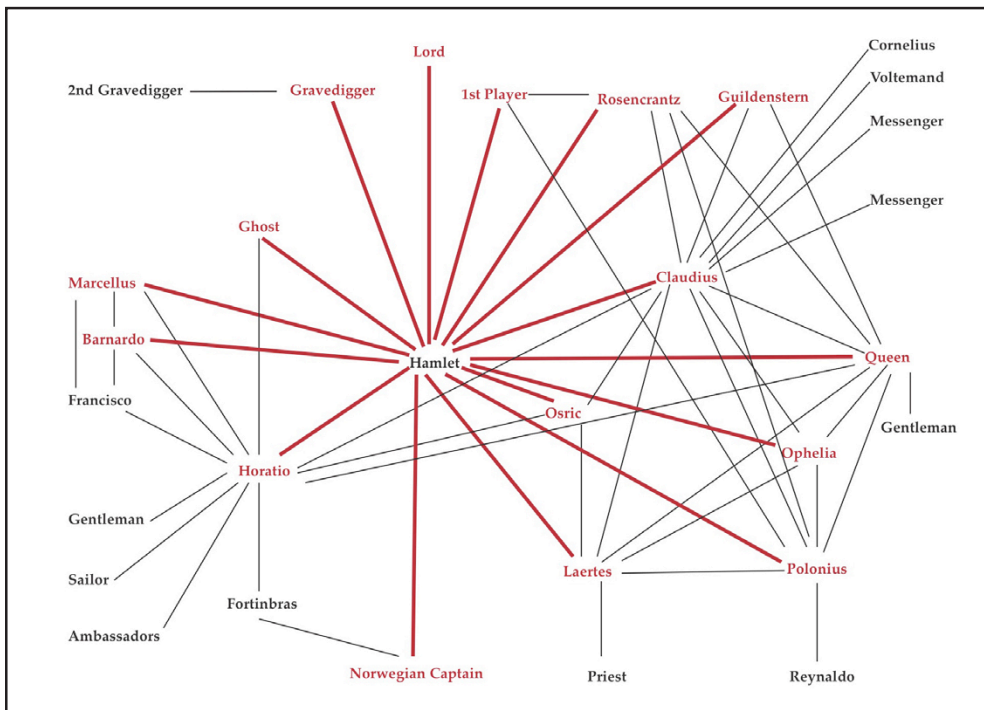


Figure 2

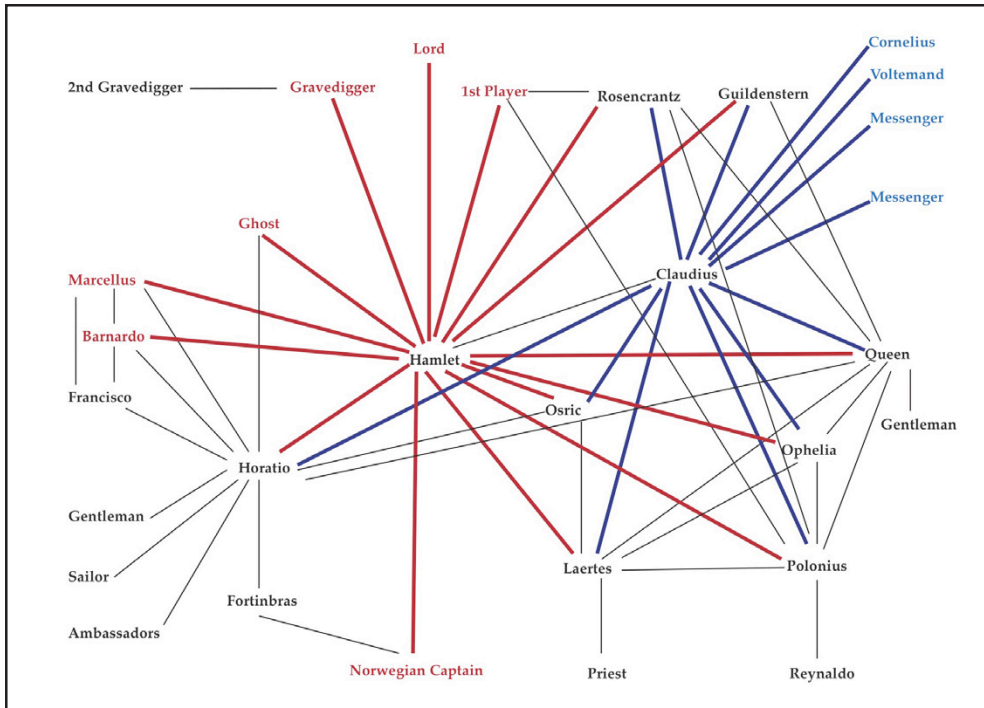


Figure 3

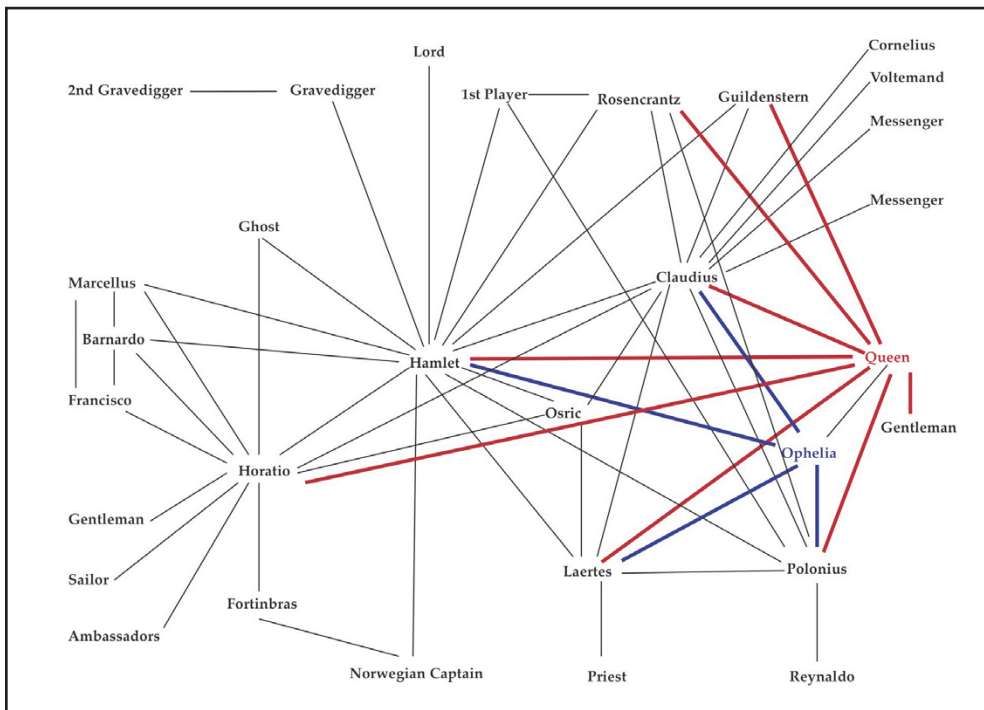


Figure 4

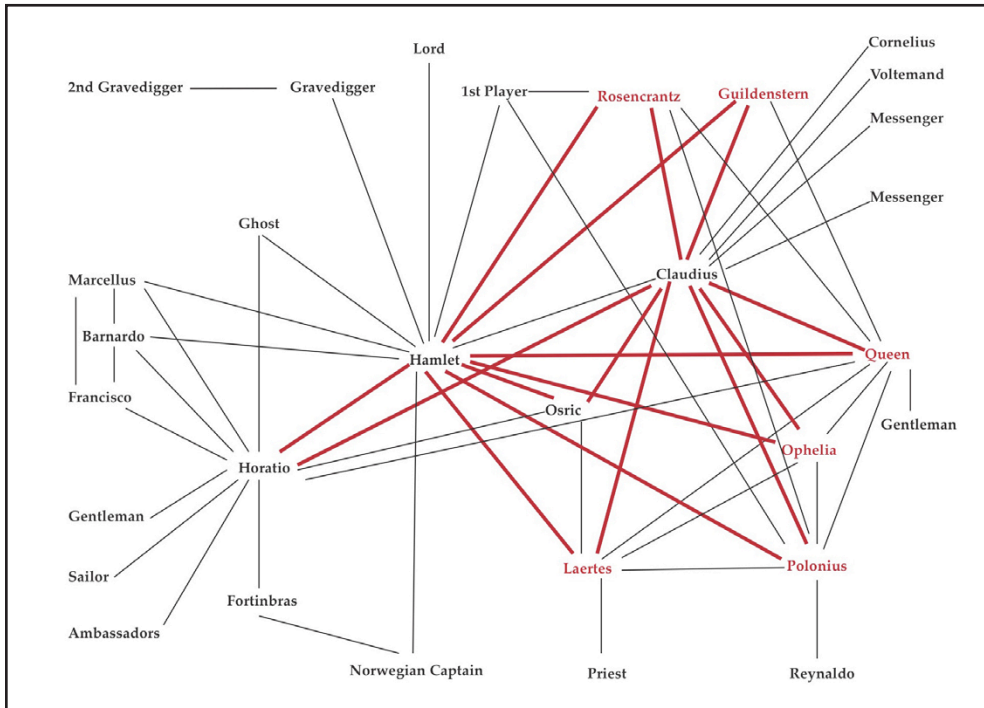


Figure 5

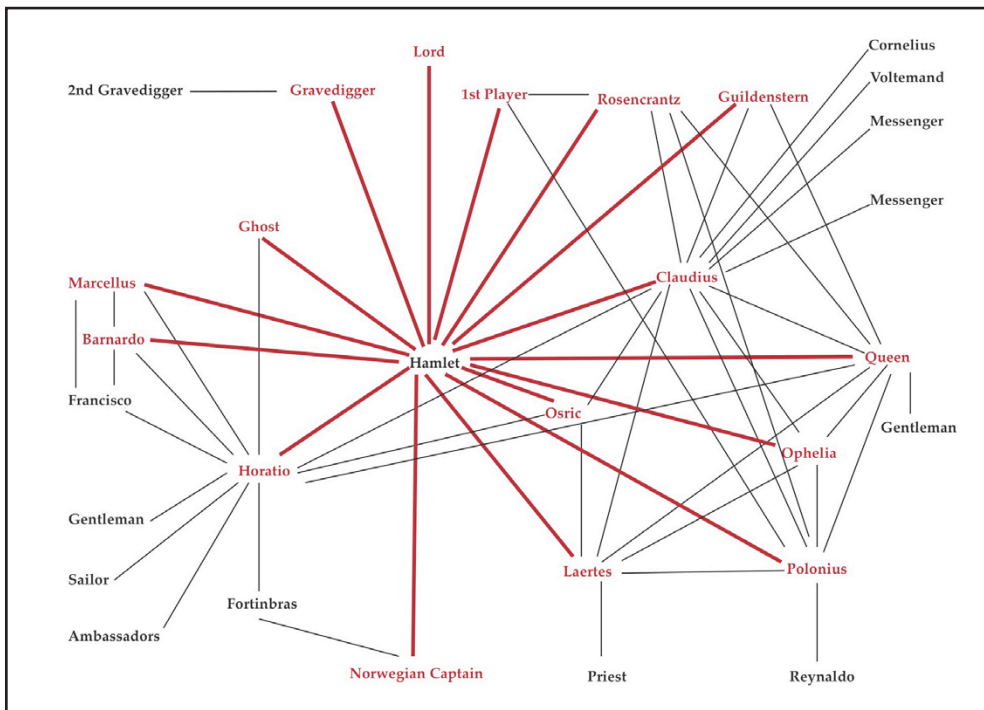


Figure 6

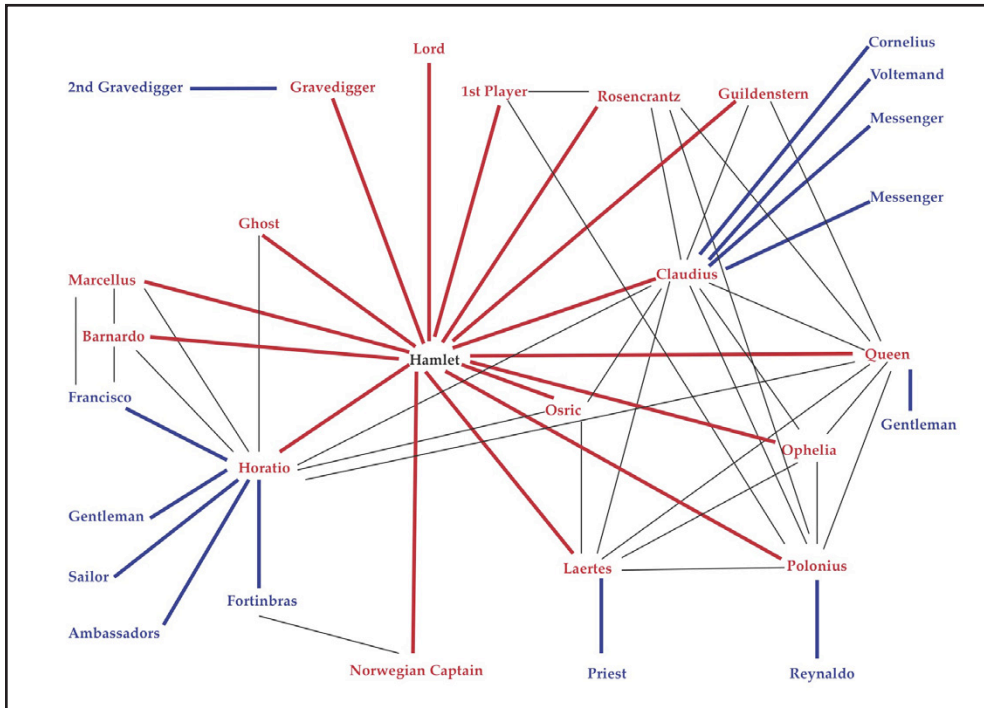


Figure 7

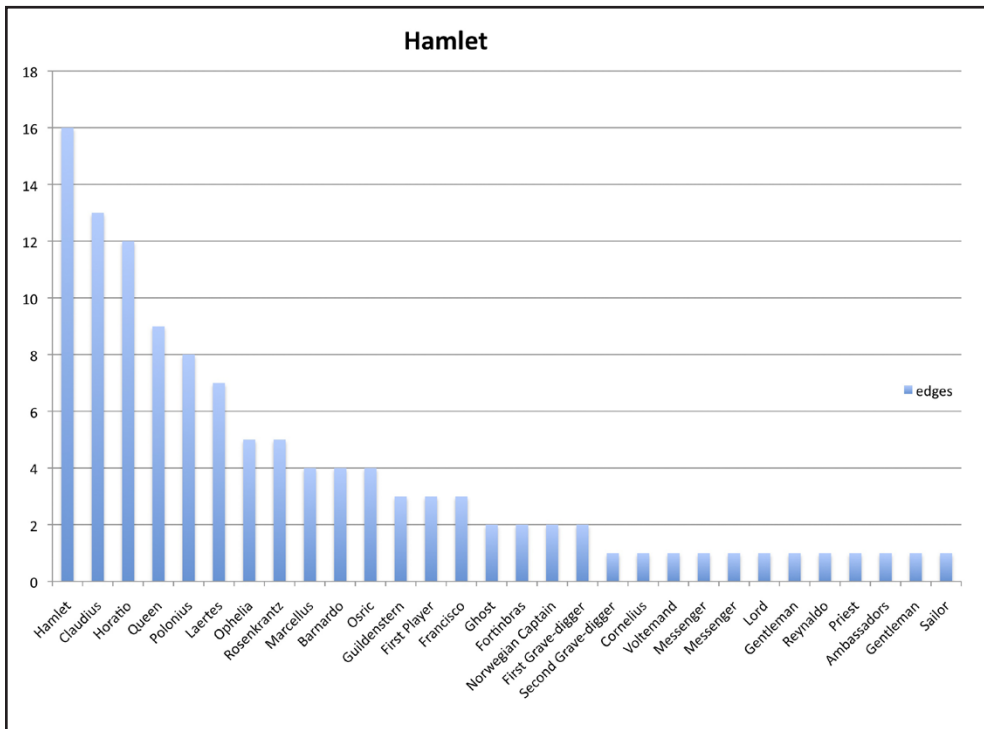


Figure 8

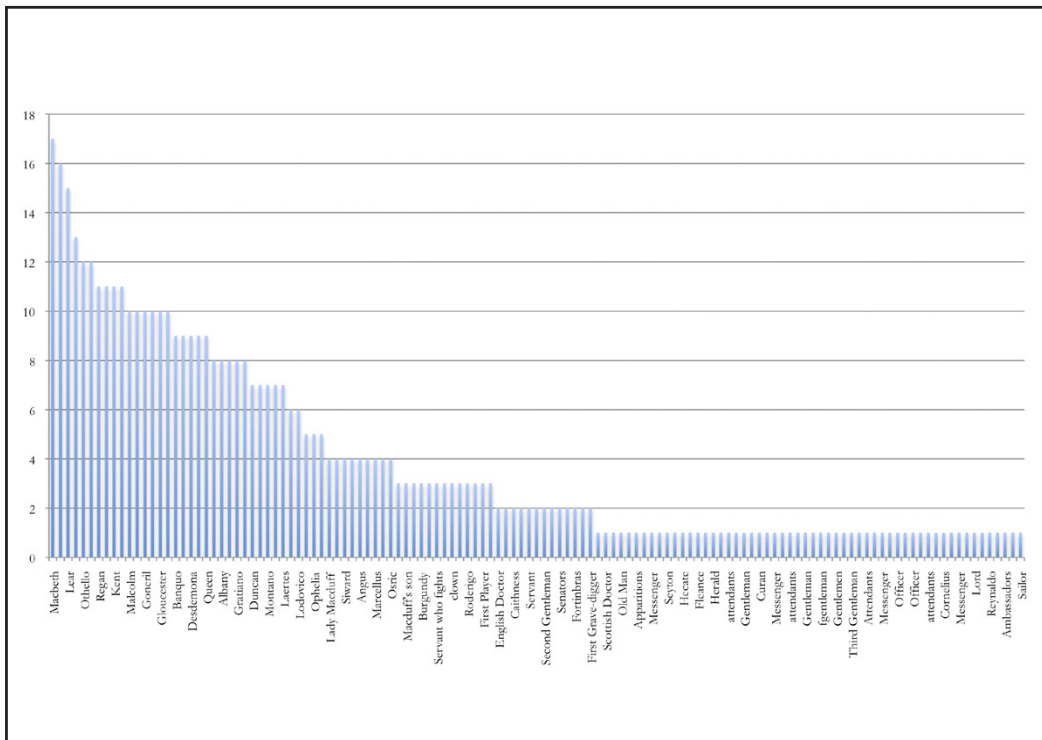


Figure 9

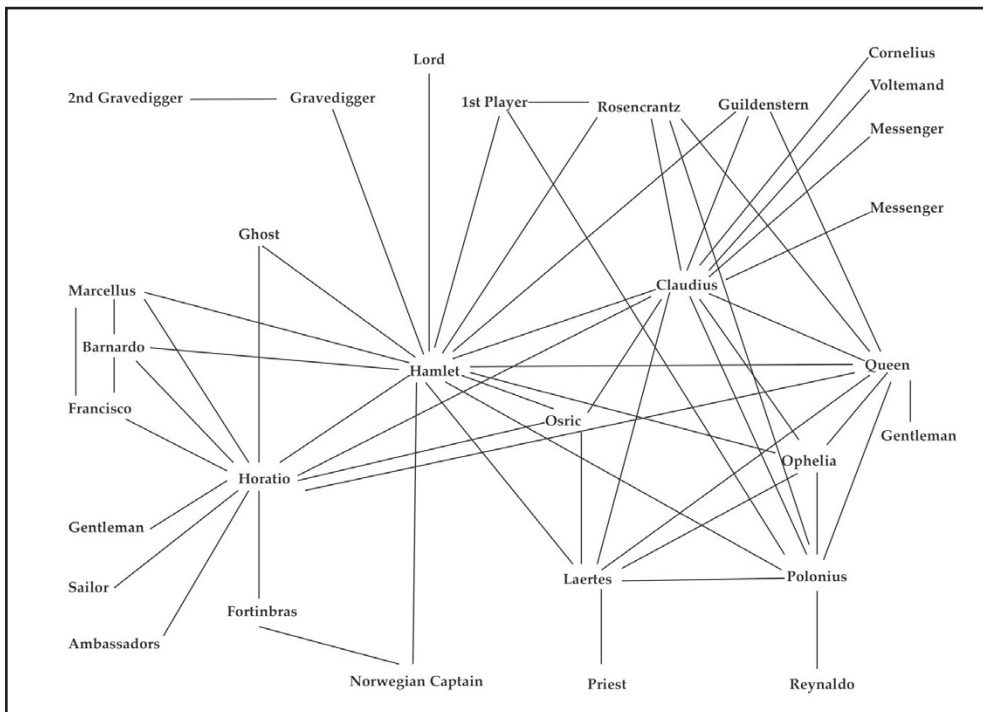


Figure 10

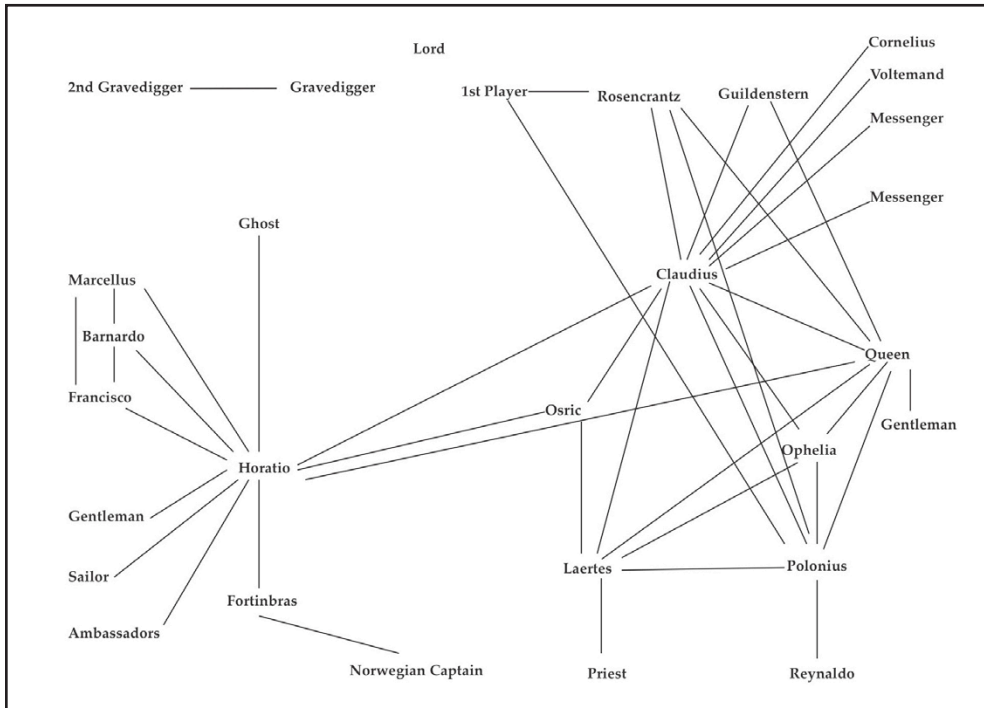


Figure 11

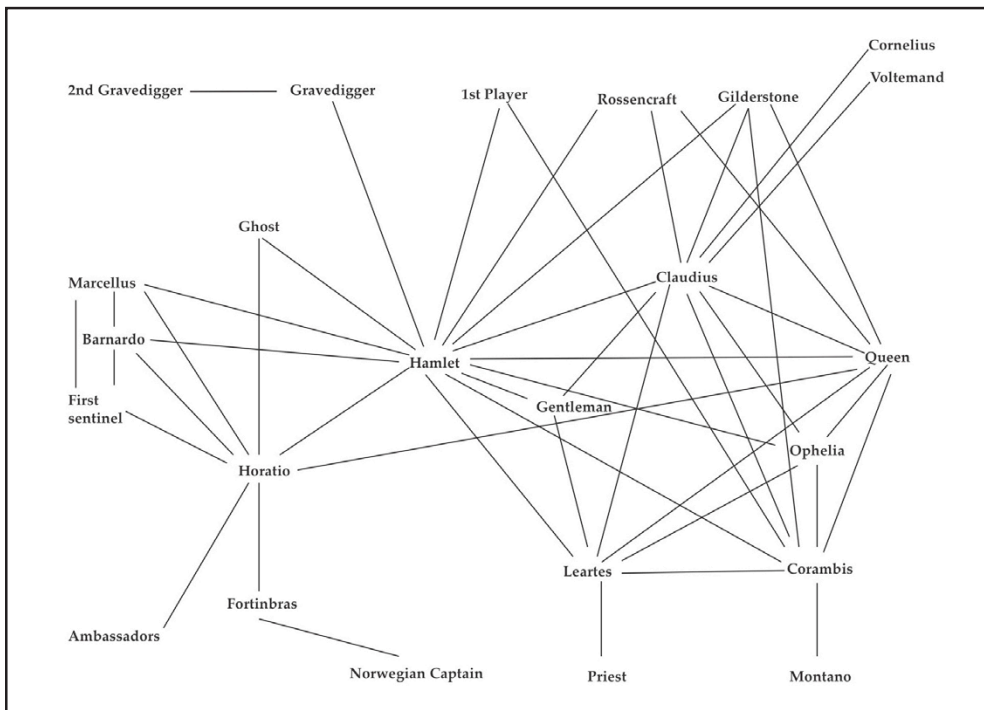


Figure 12

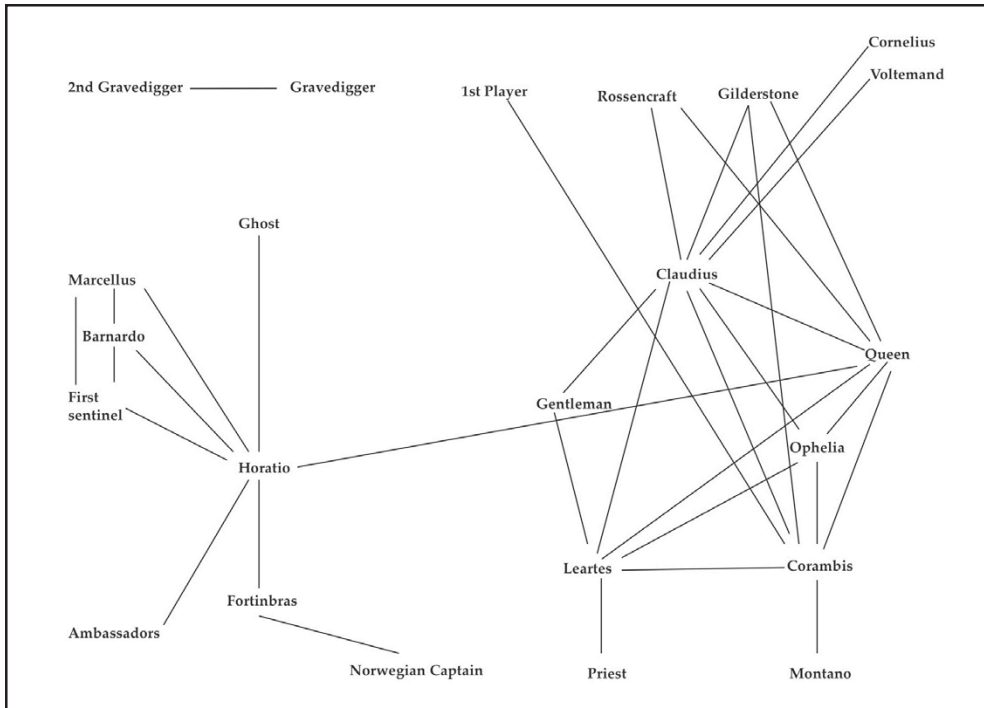


Figure 13

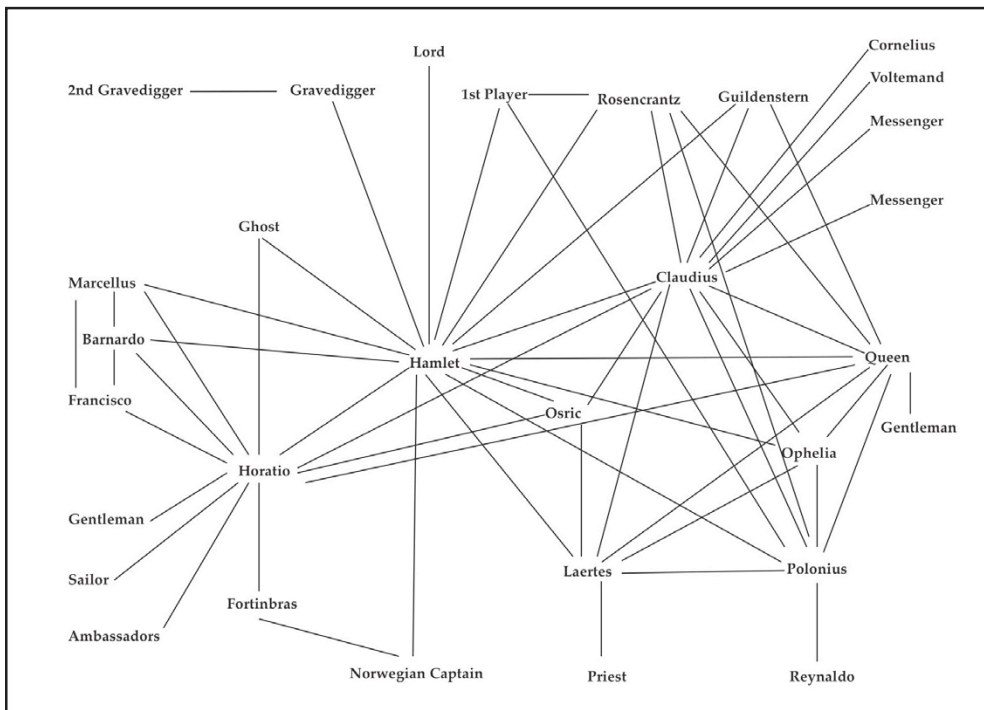


Figure 14

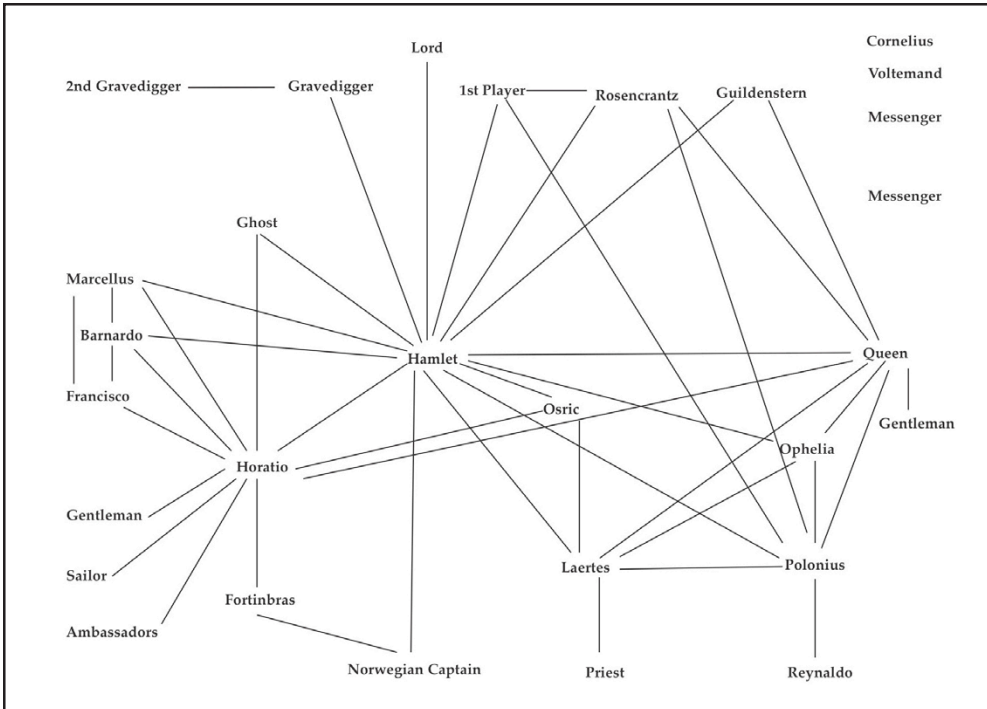


Figure 15

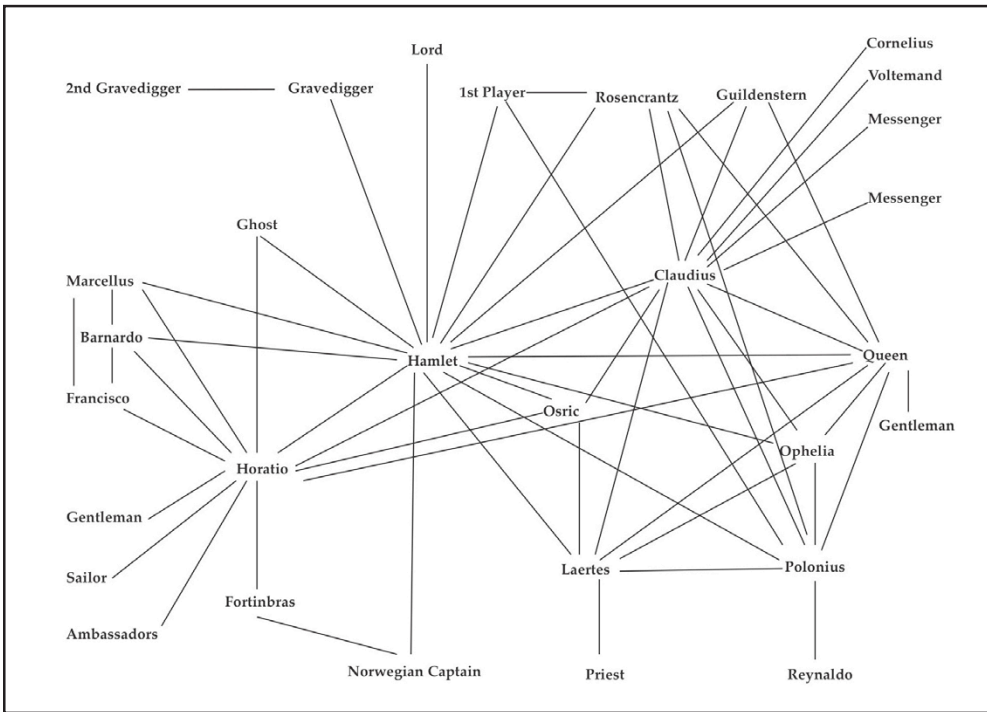


Figure 16

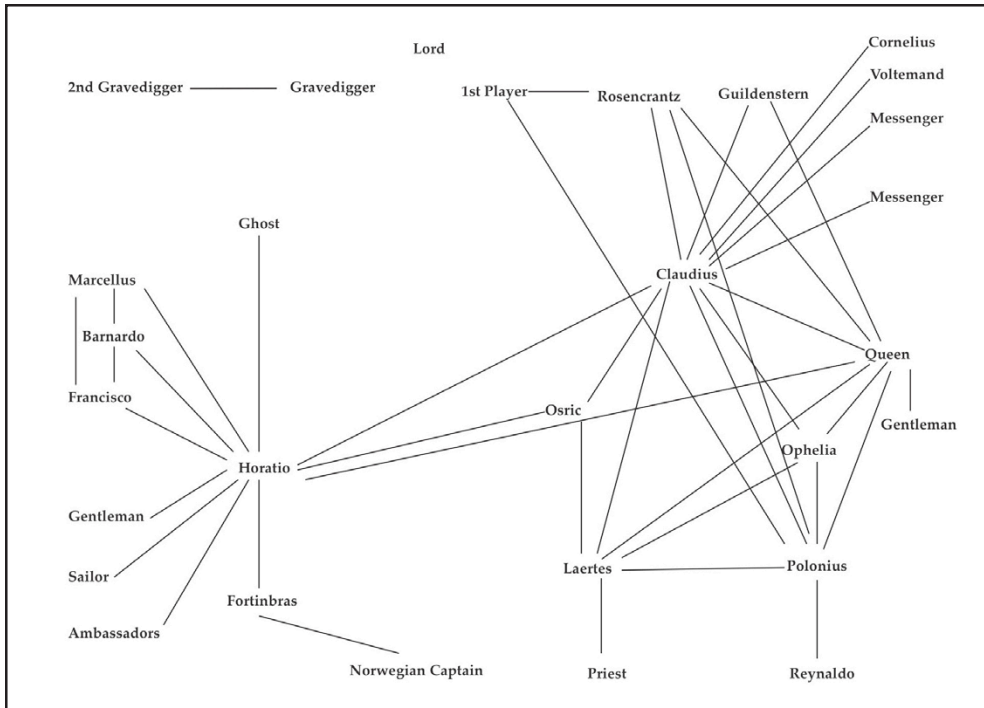


Figure 17

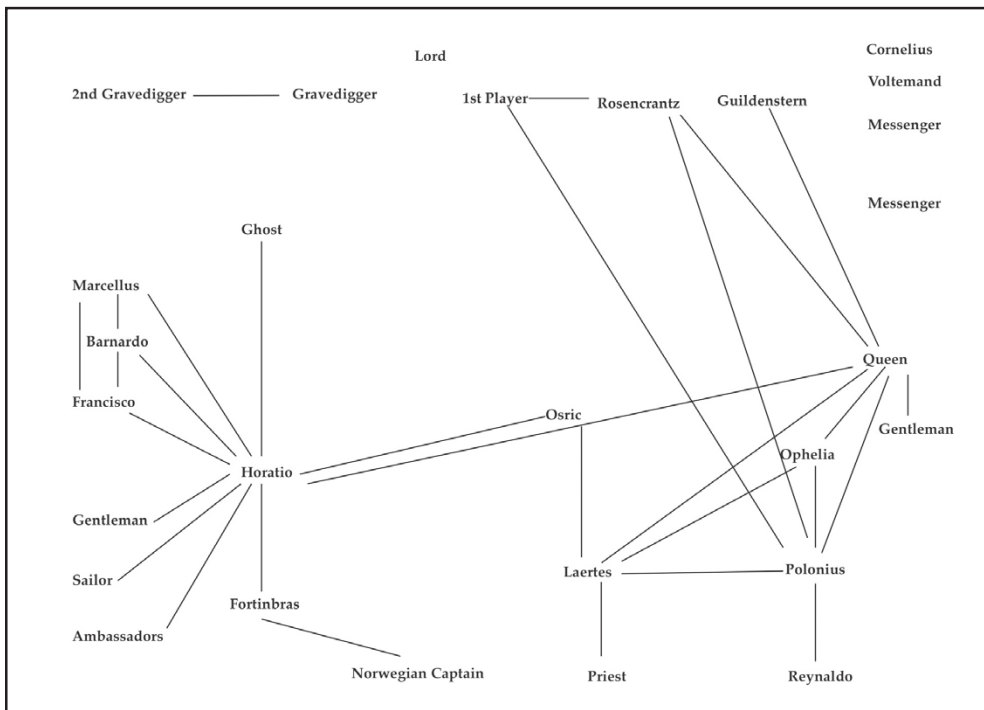


Figure 18

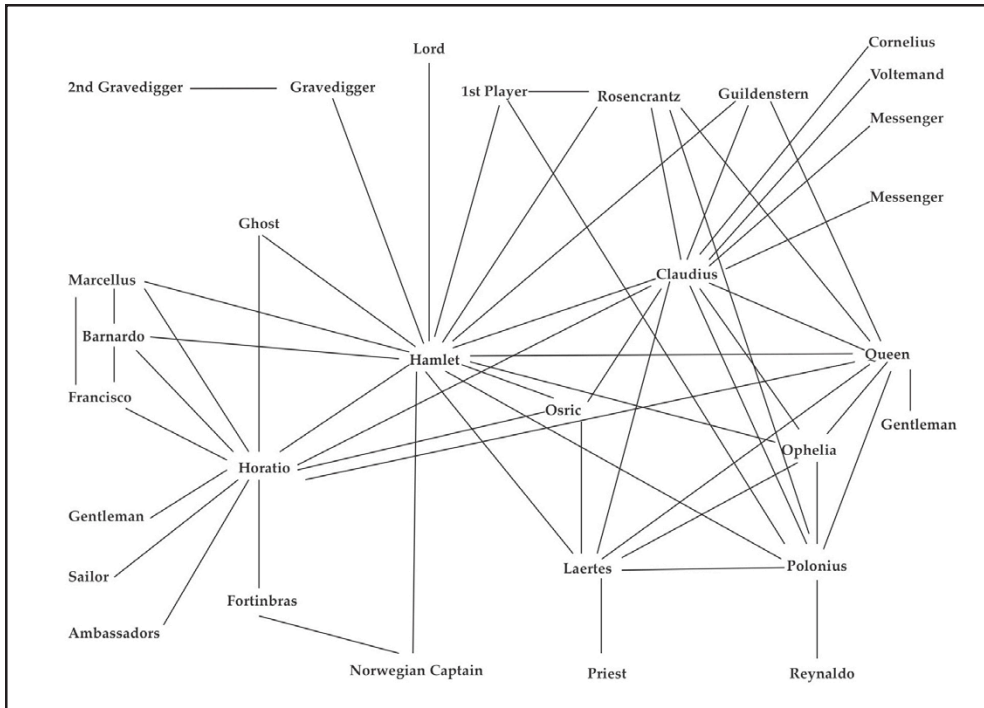


Figure 19

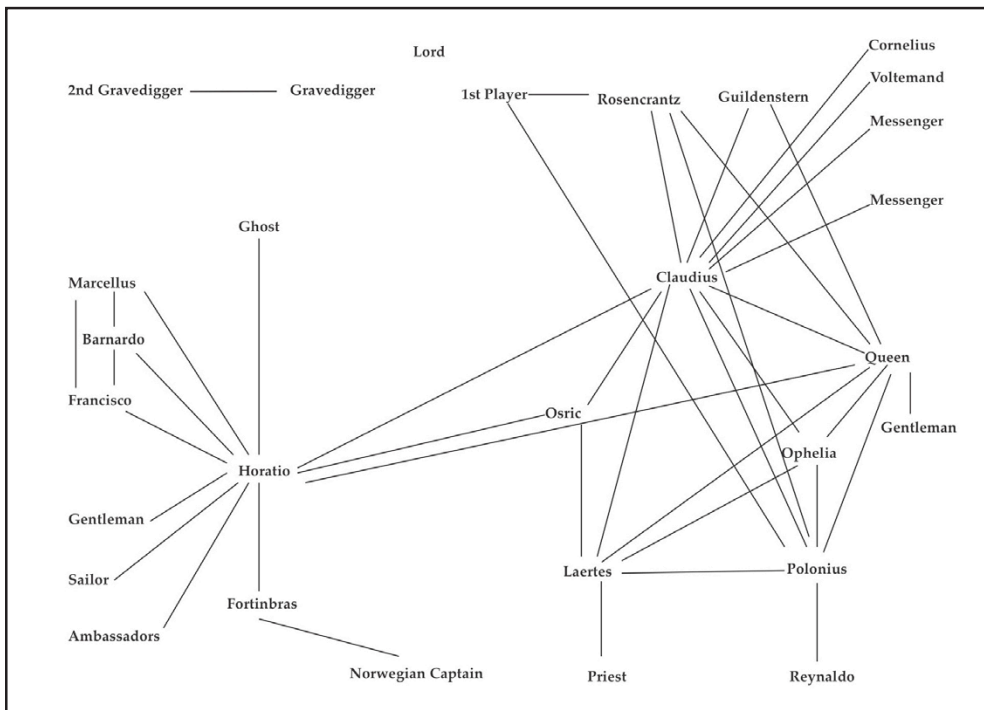


Figure 20

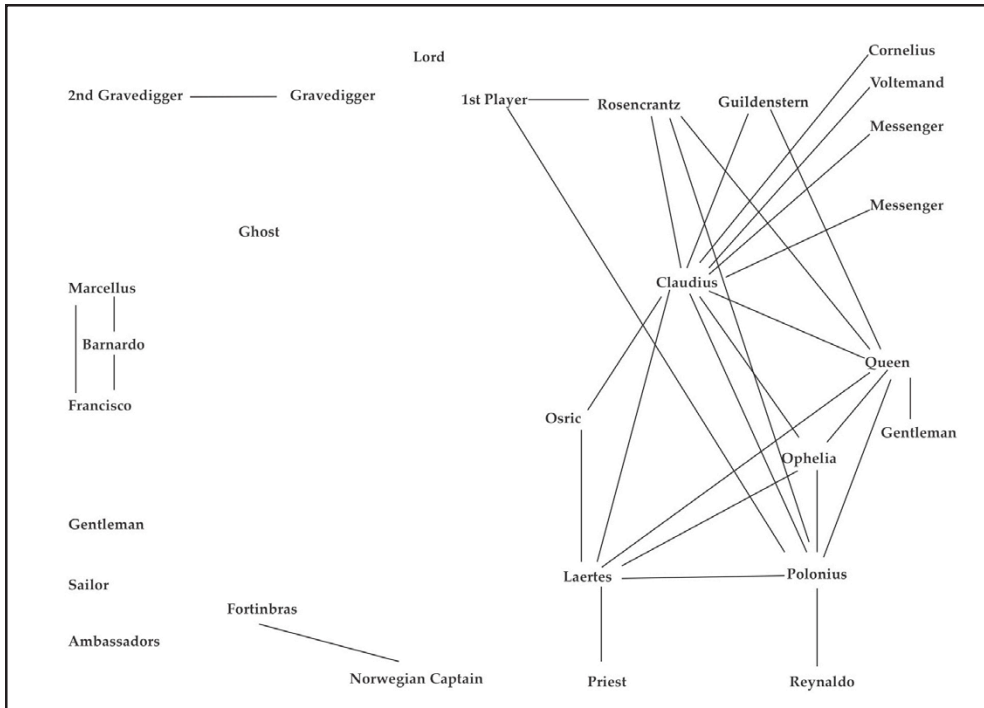


Figure 21

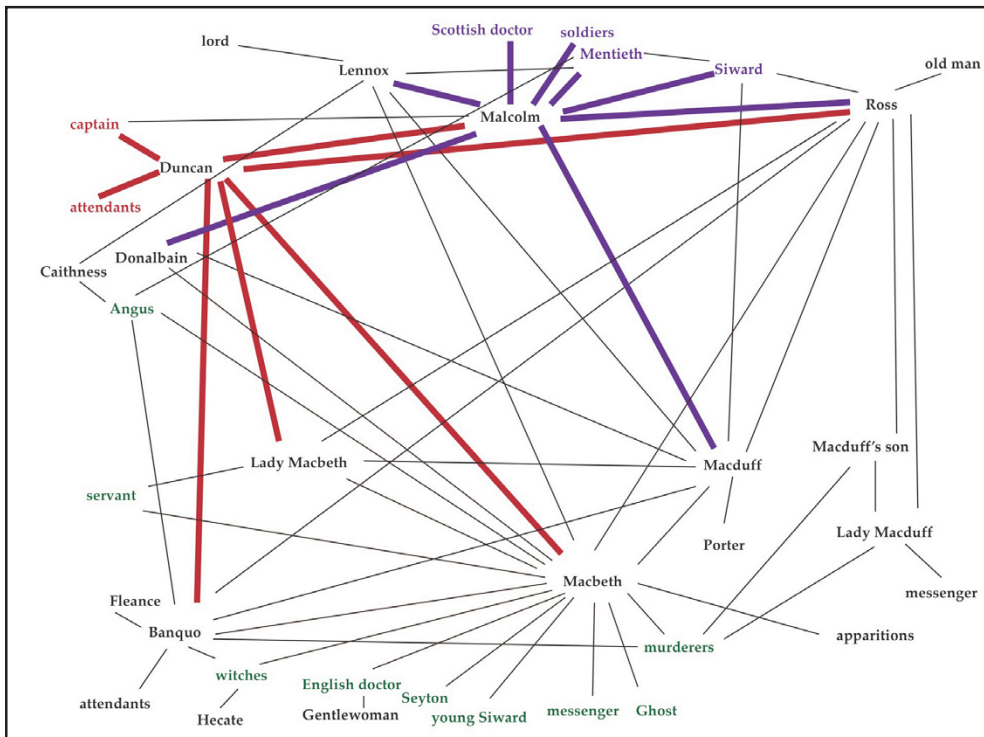


Figure 22

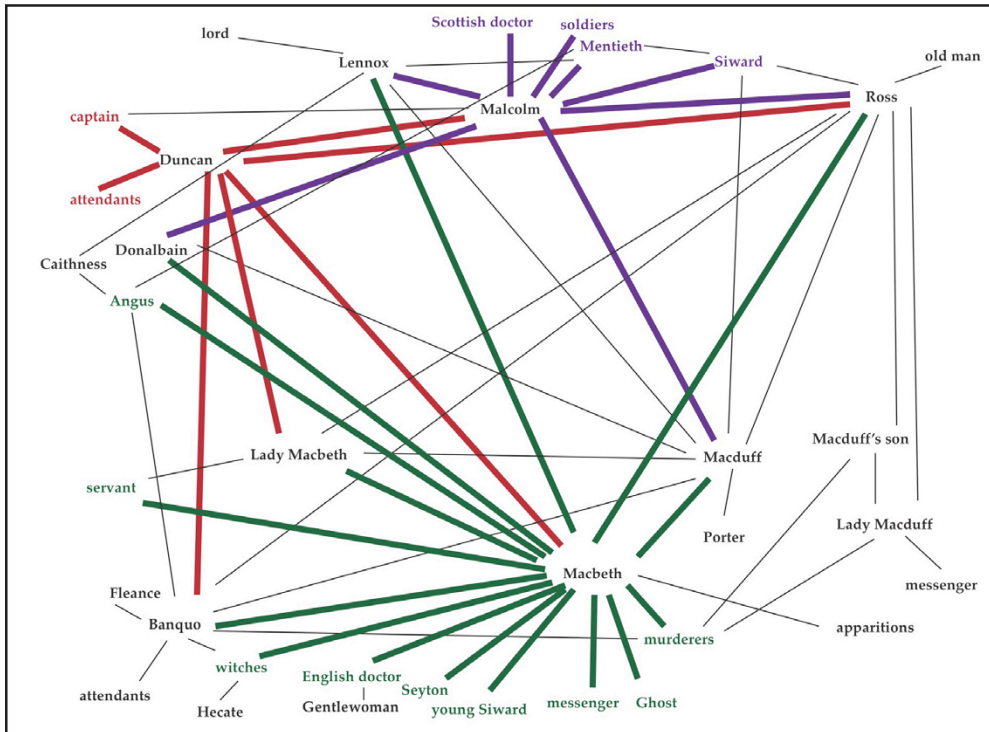


Figure 23

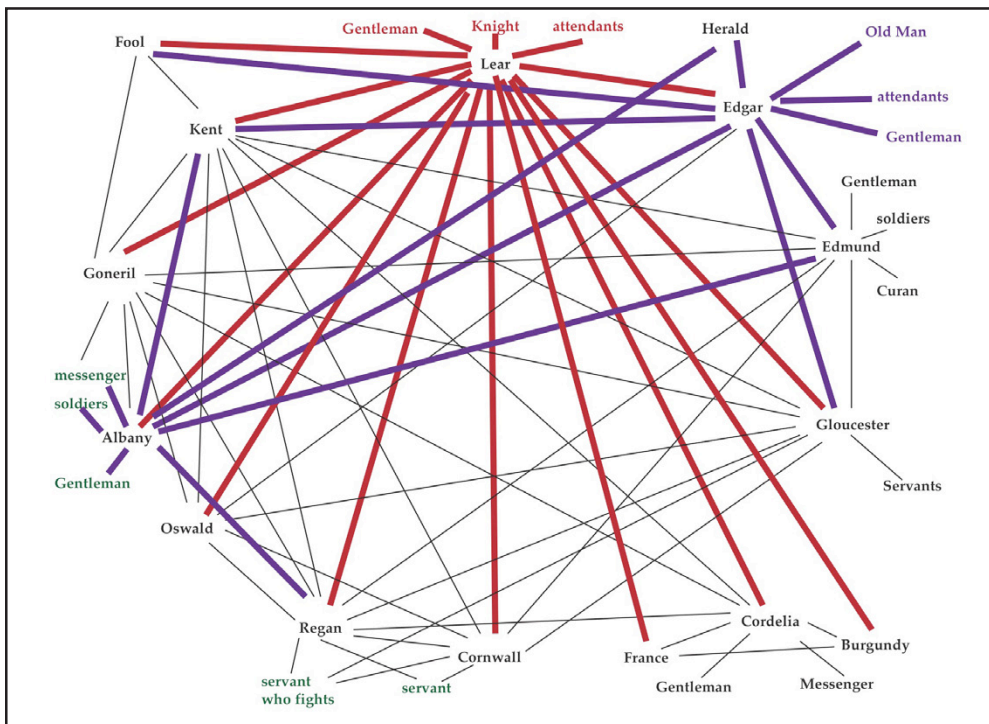


Figure 24

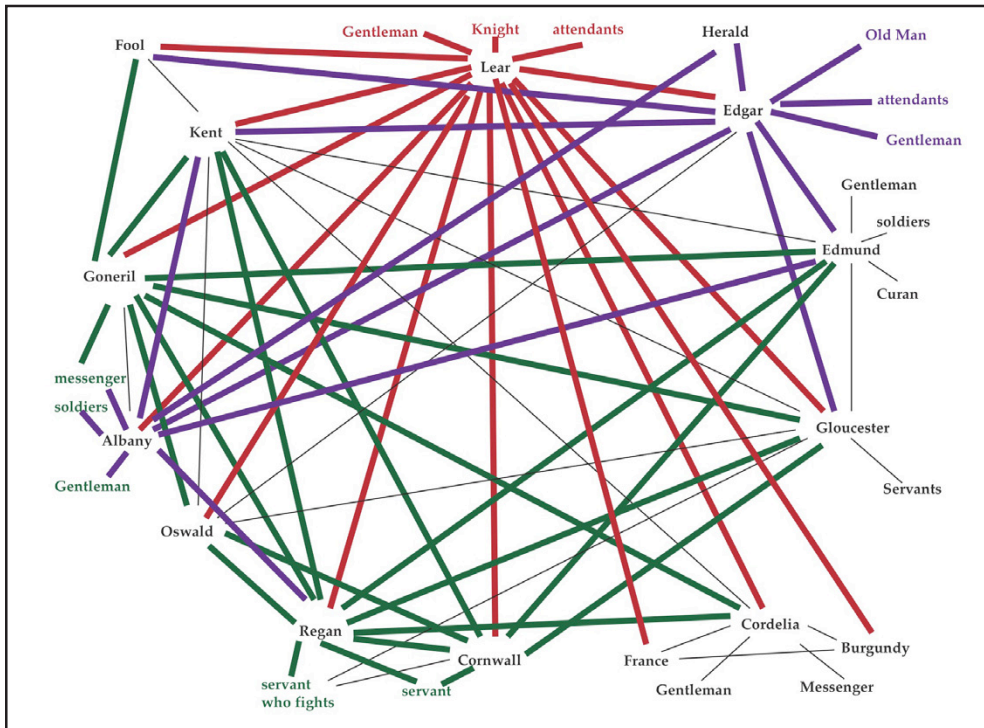


Figure 25

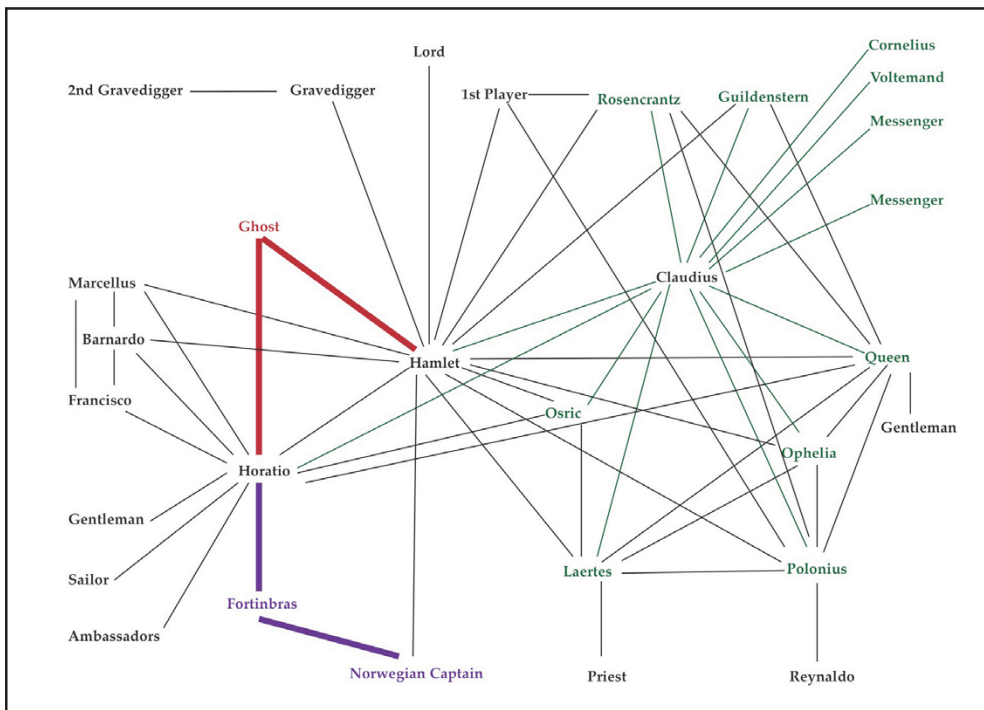


Figure 26

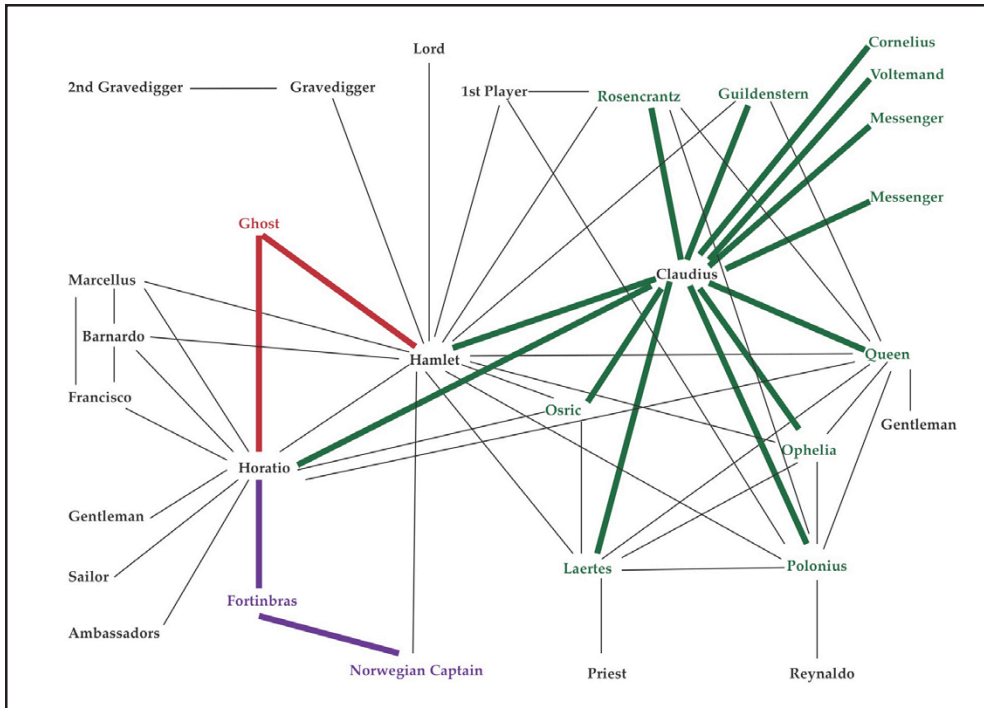


Figure 27

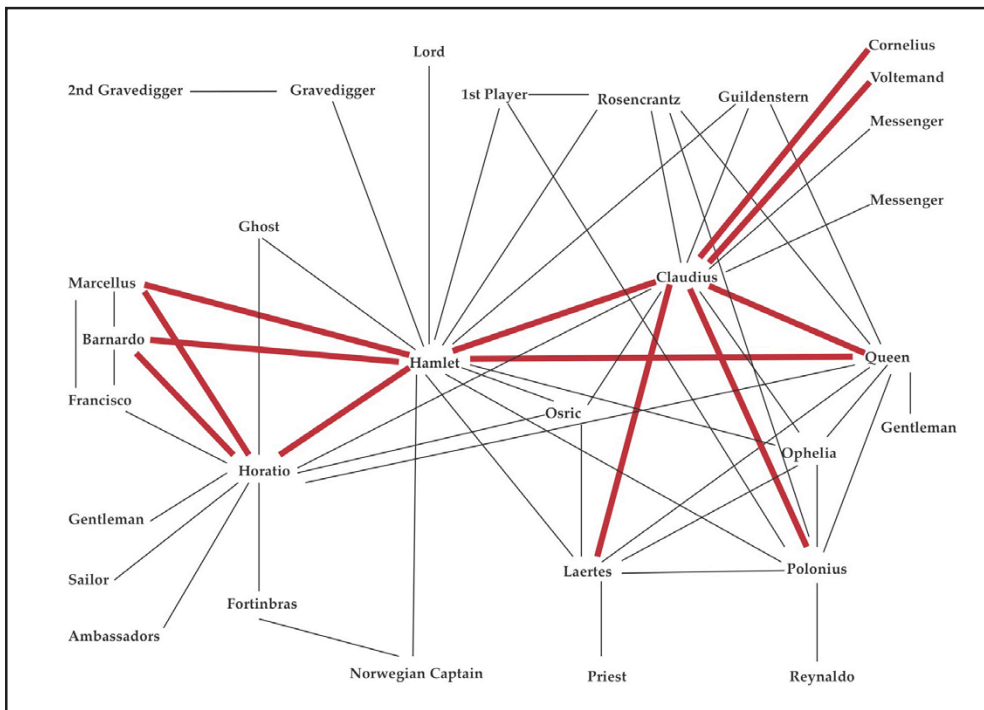


Figure 28

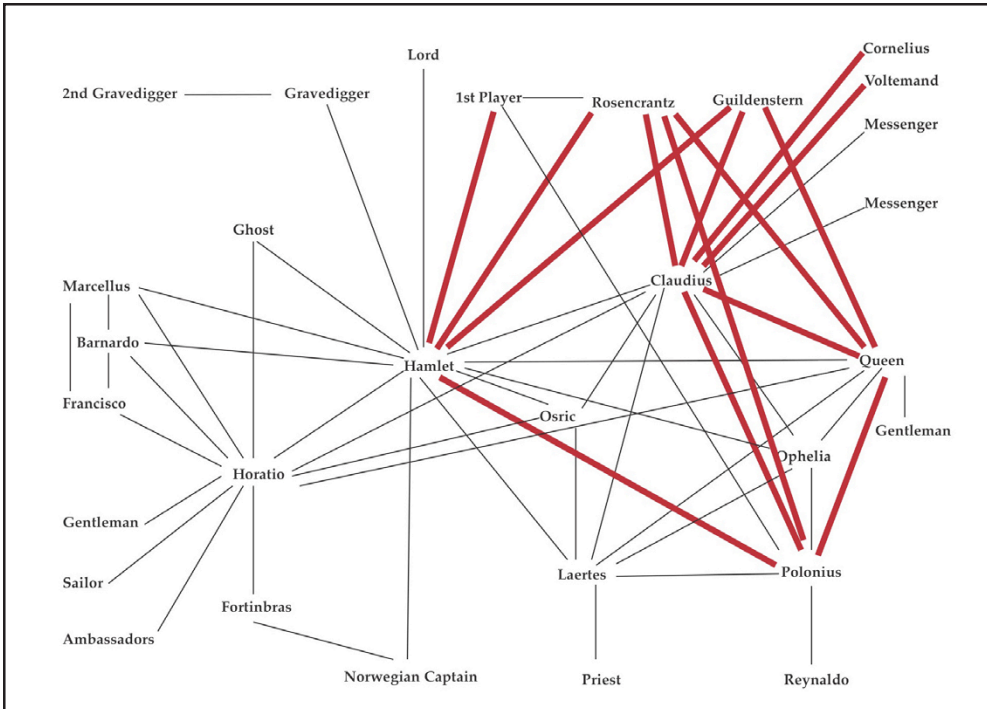


Figure 29

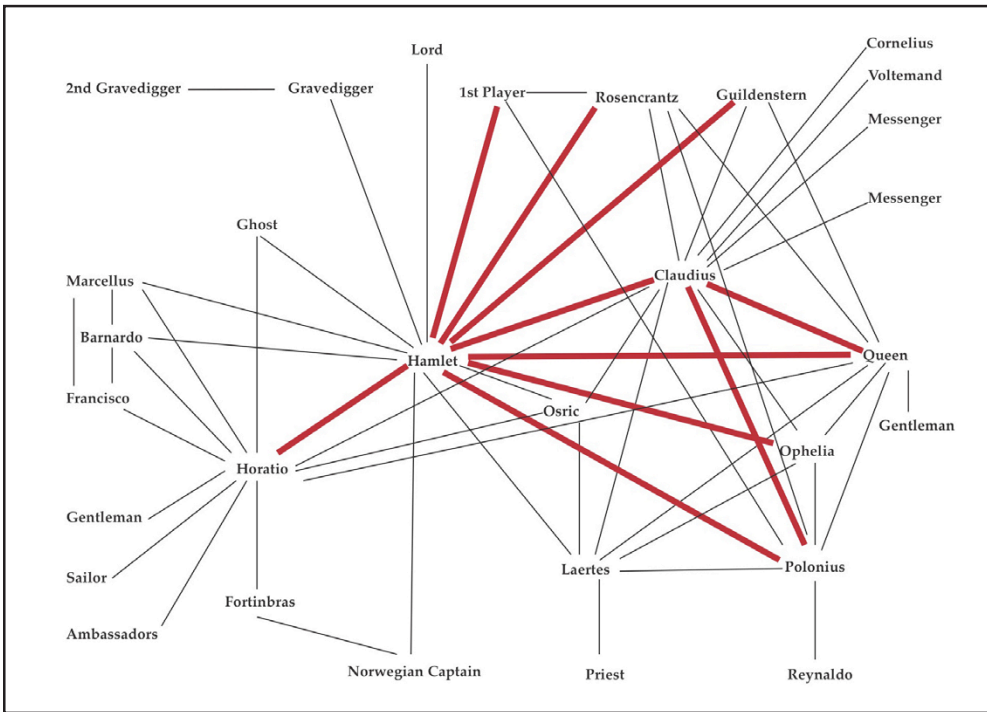


Figure 30

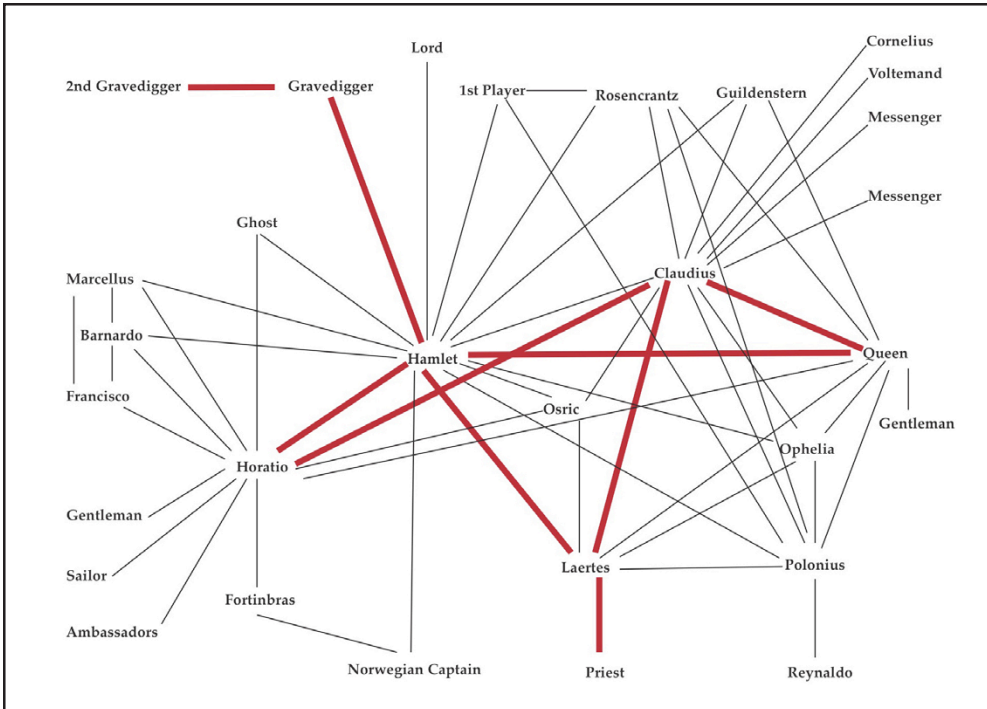


Figure 31

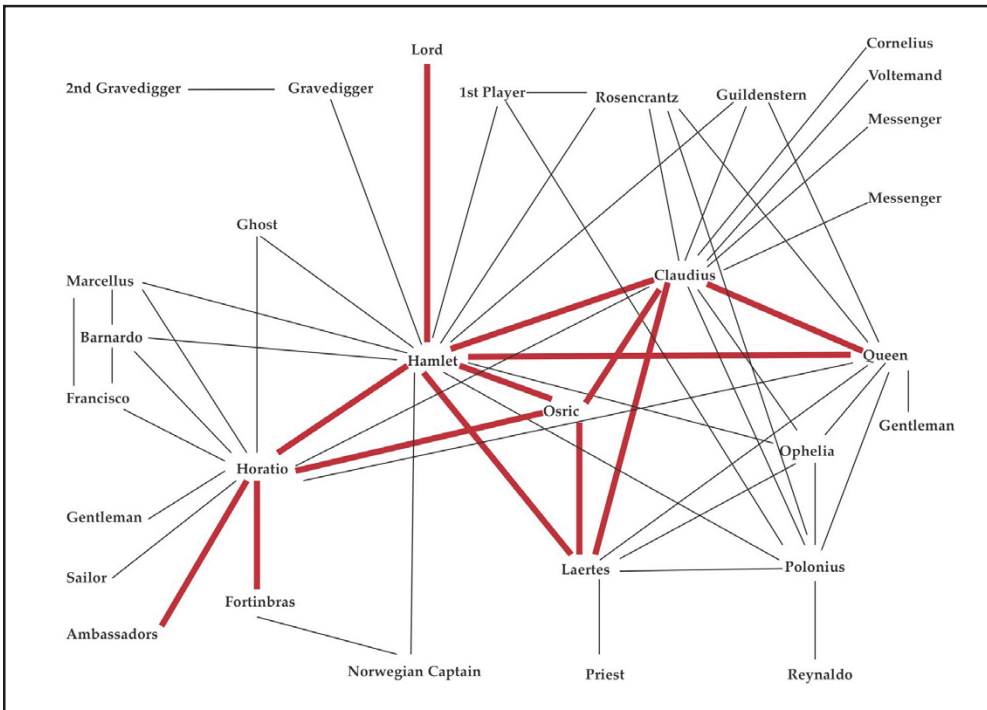


Figure 32

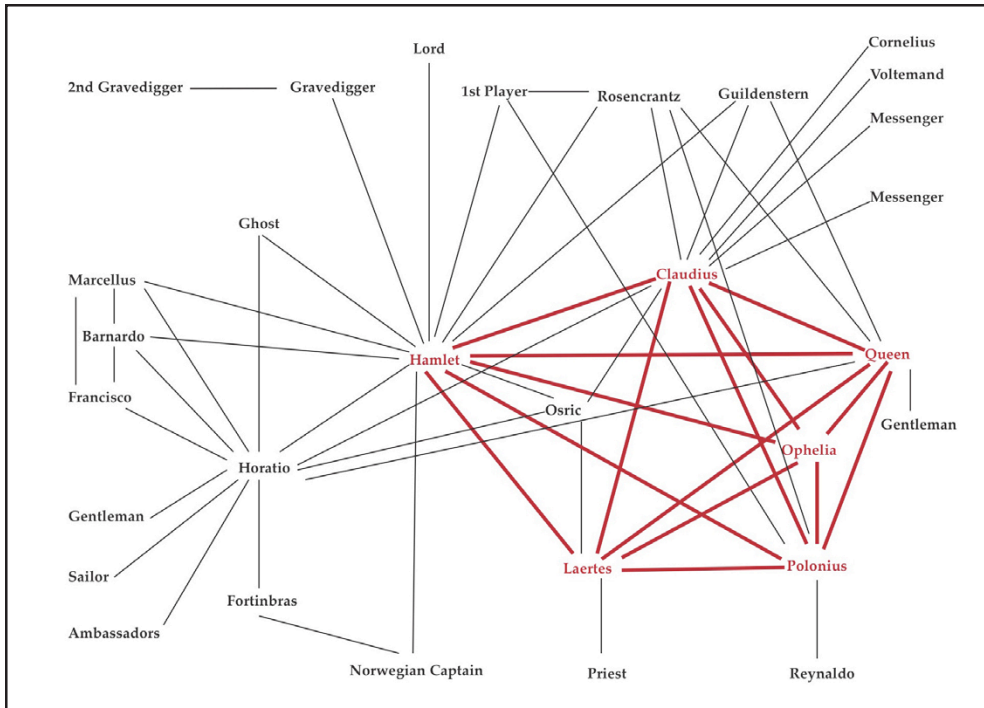


Figure 33

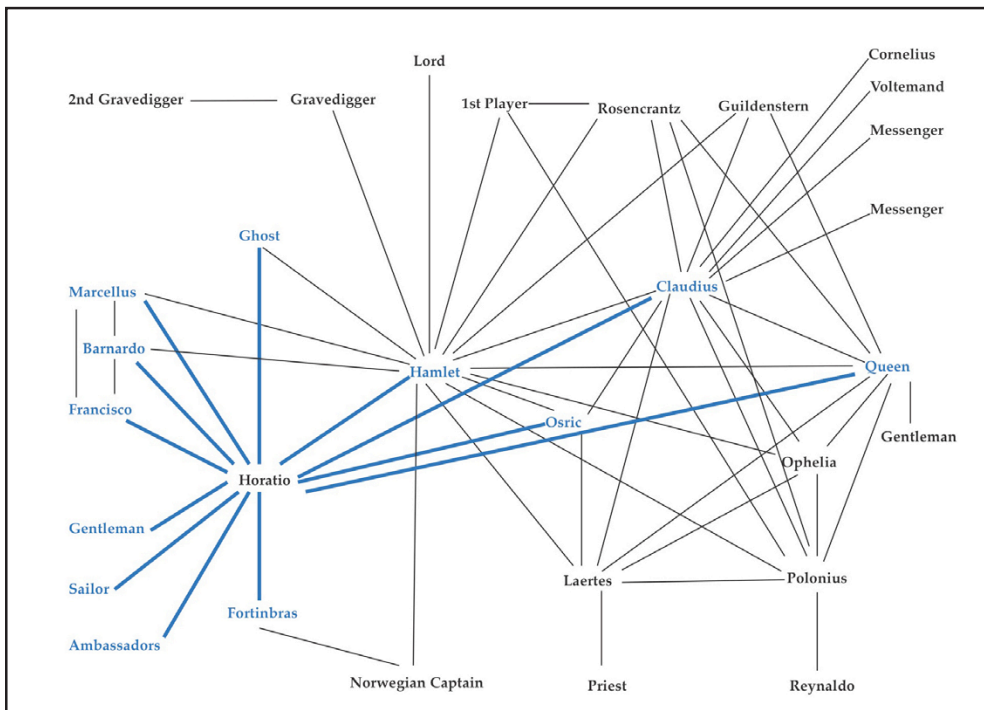


Figure 34

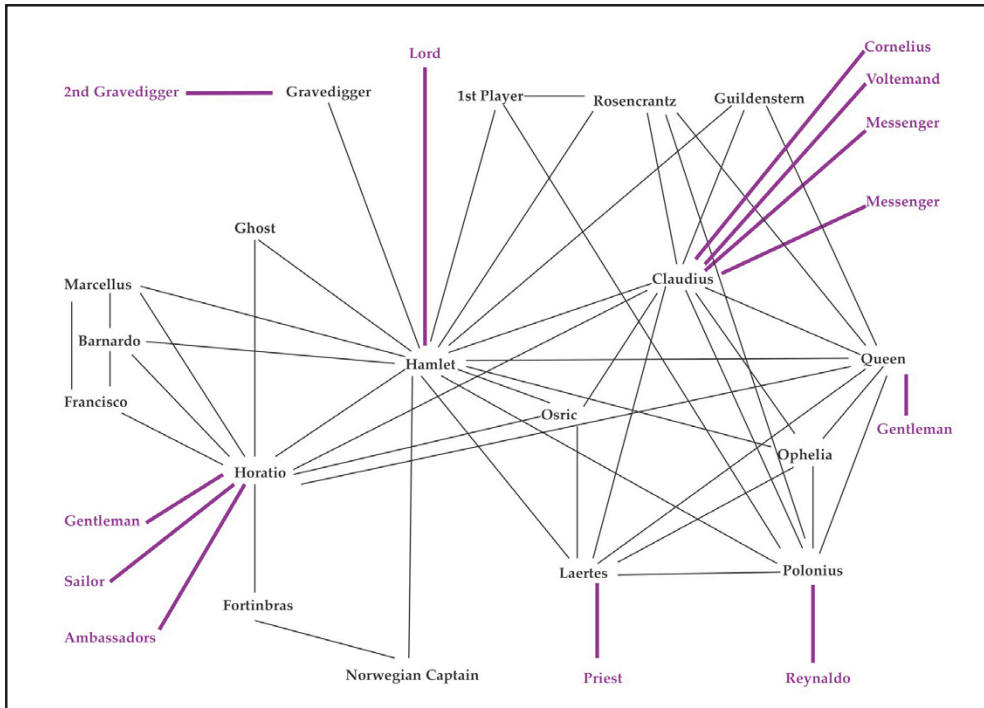


Figure 35

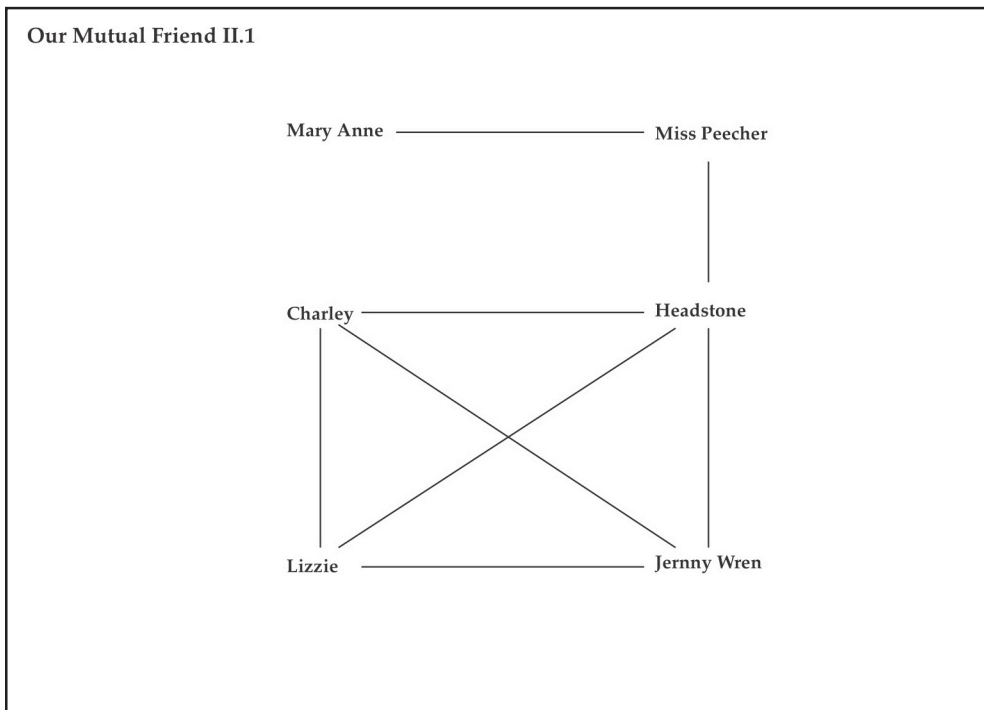


Figure 36

Our Mutual Friend II.2

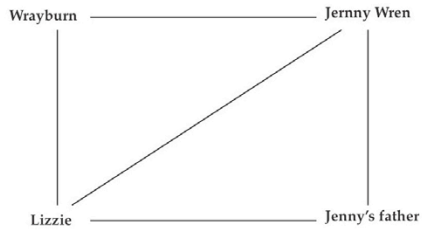


Figure 37

Our Mutual Friend II.4

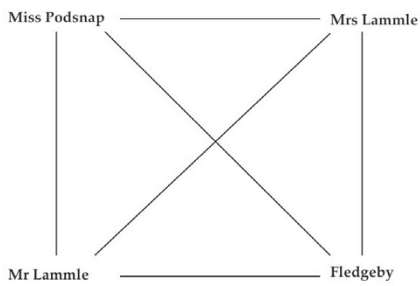


Figure 38

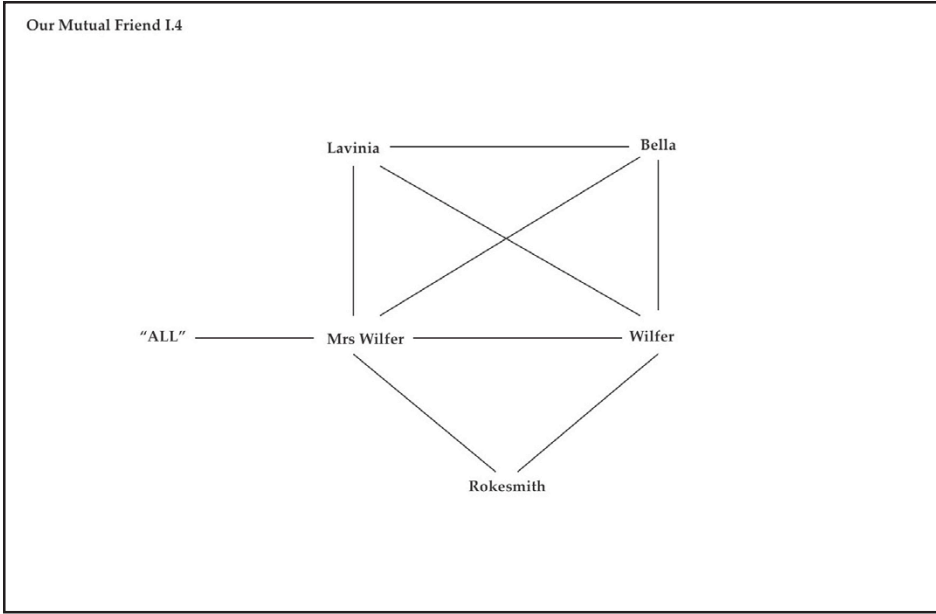


Figure 39

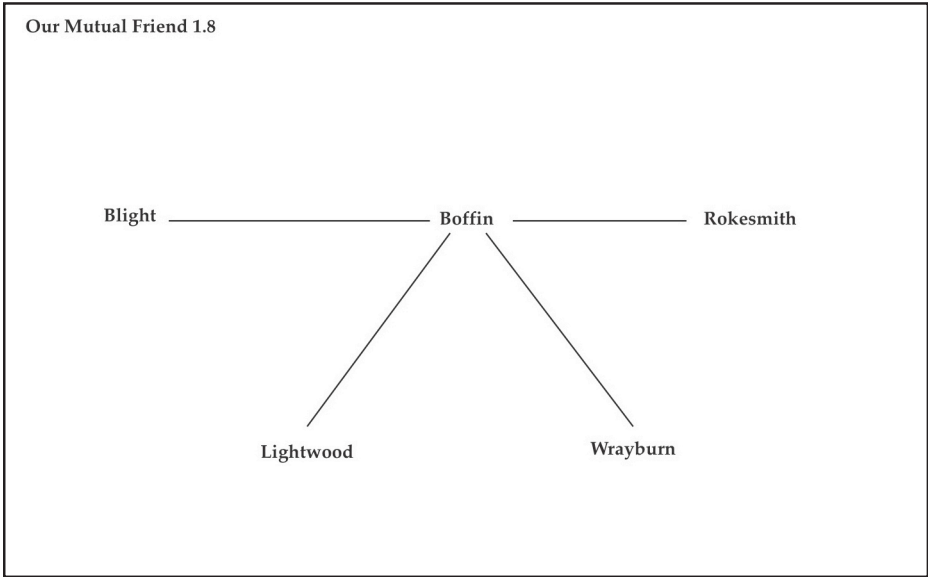


Figure 40

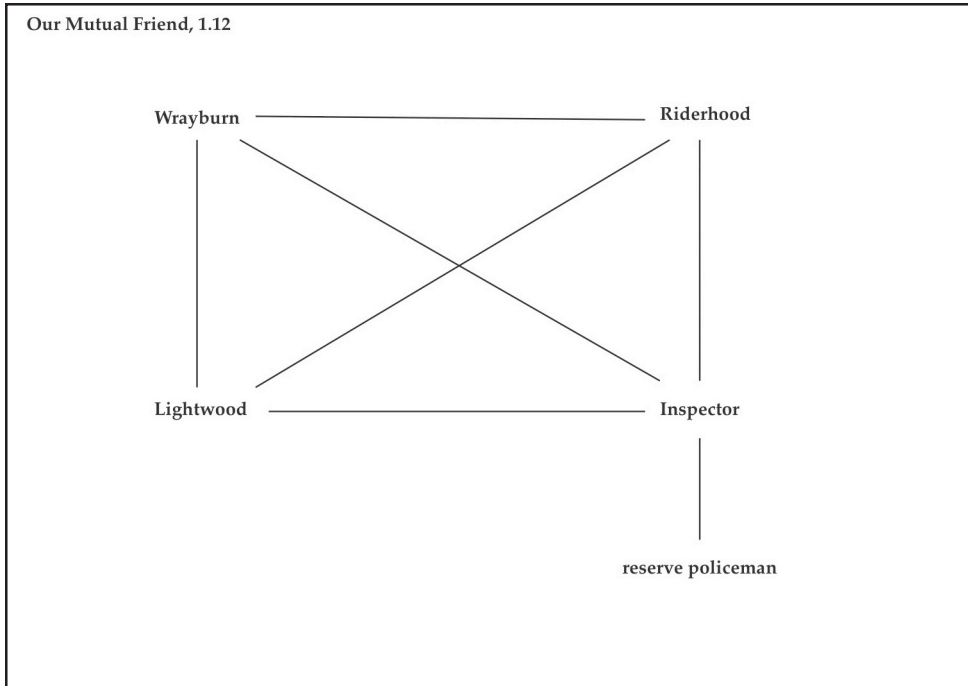


Figure 41

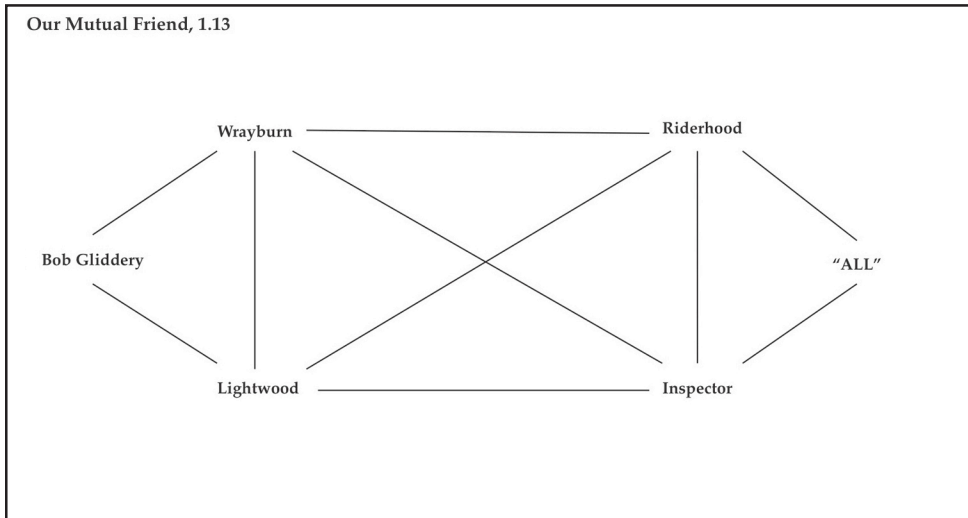


Figure 42

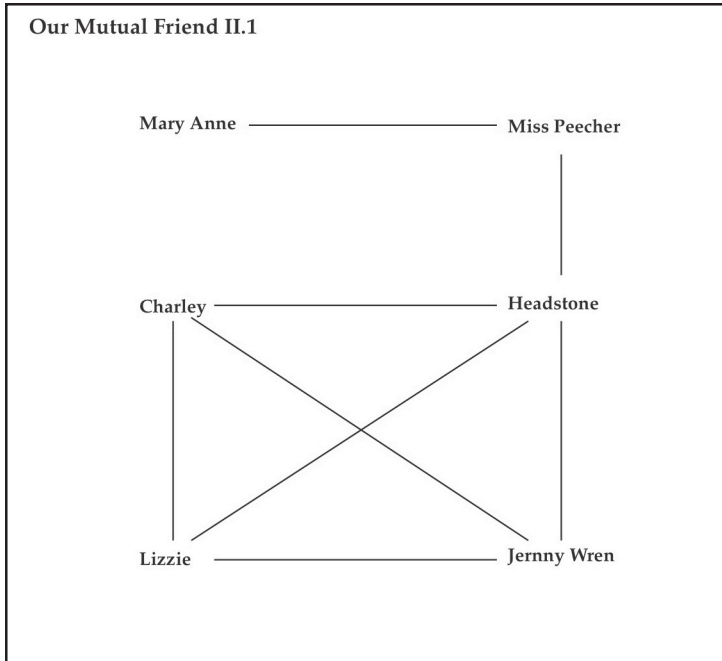


Figure 43

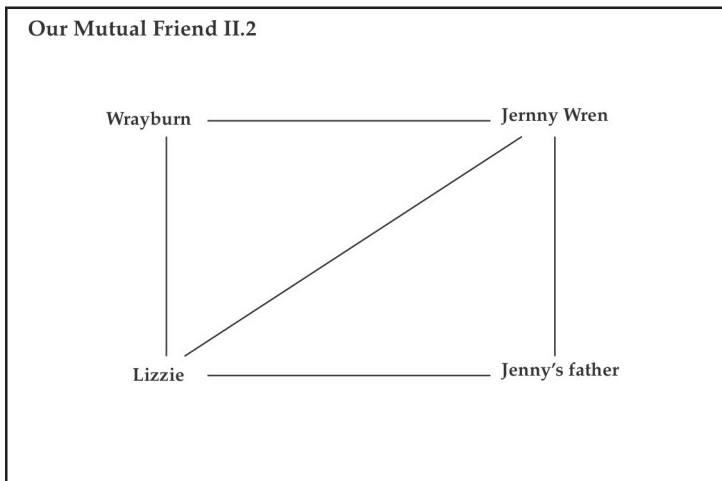


Figure 44

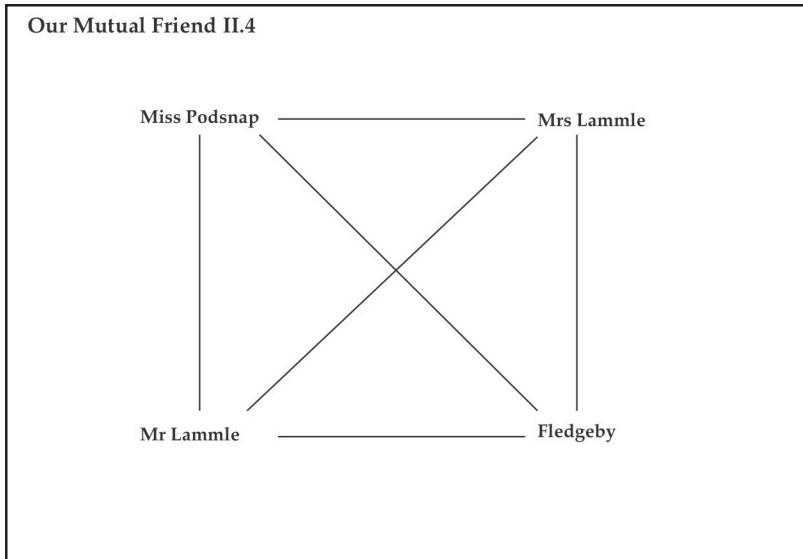


Figure 45

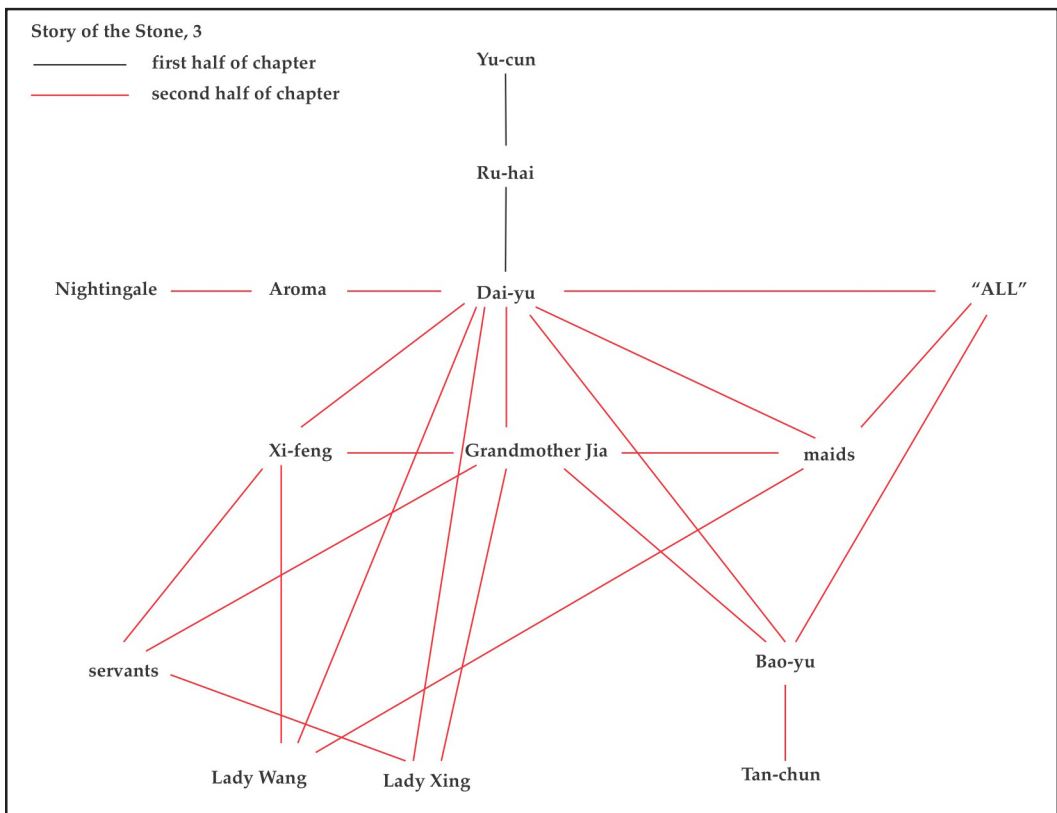


Figure 46

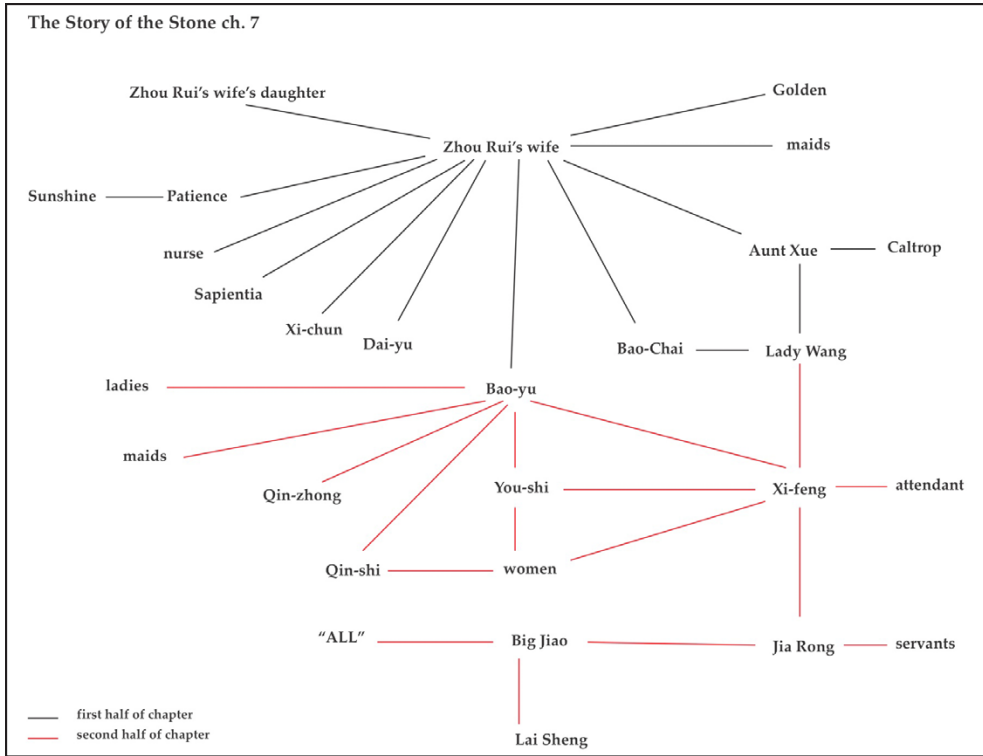


Figure 47

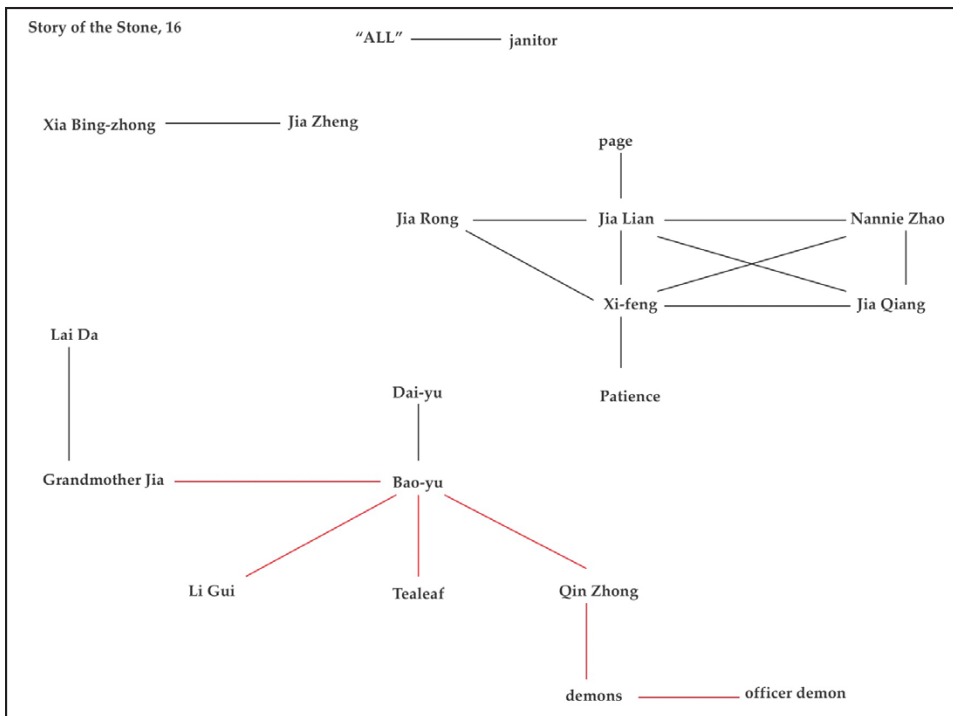


Figure 48

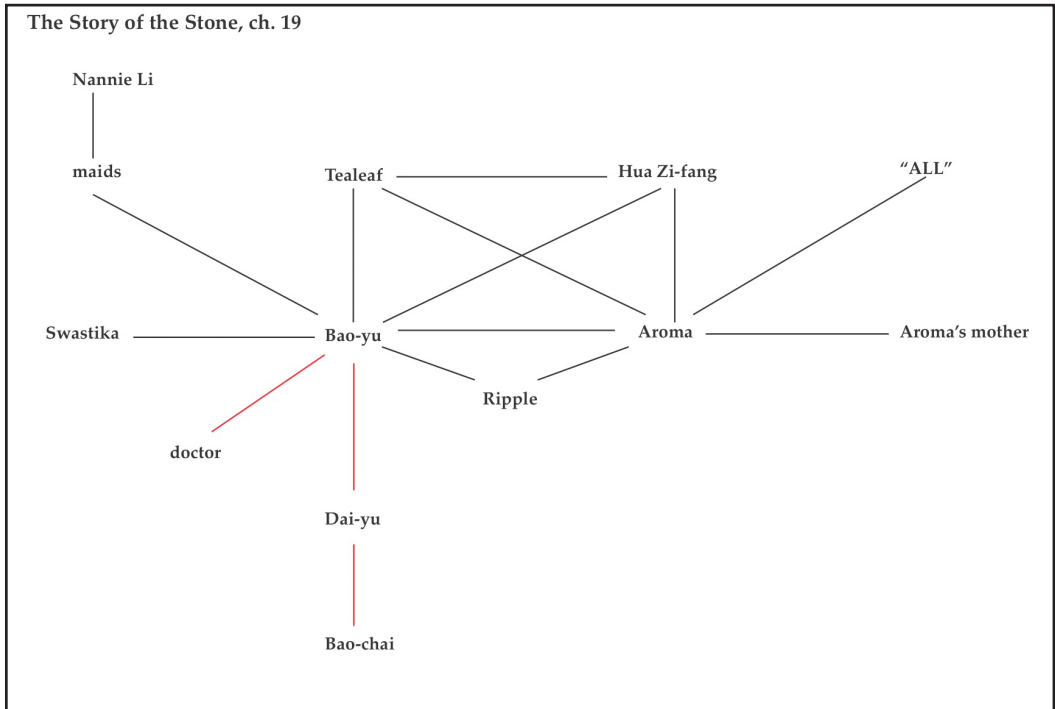


Figure 49

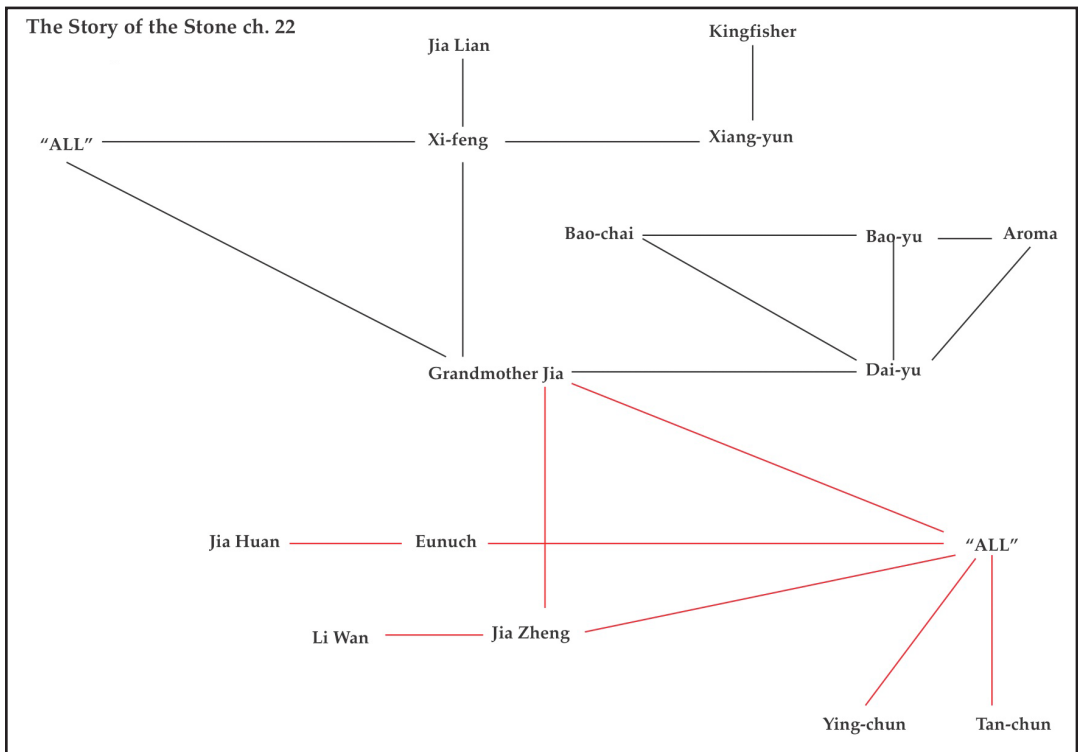


Figure 50

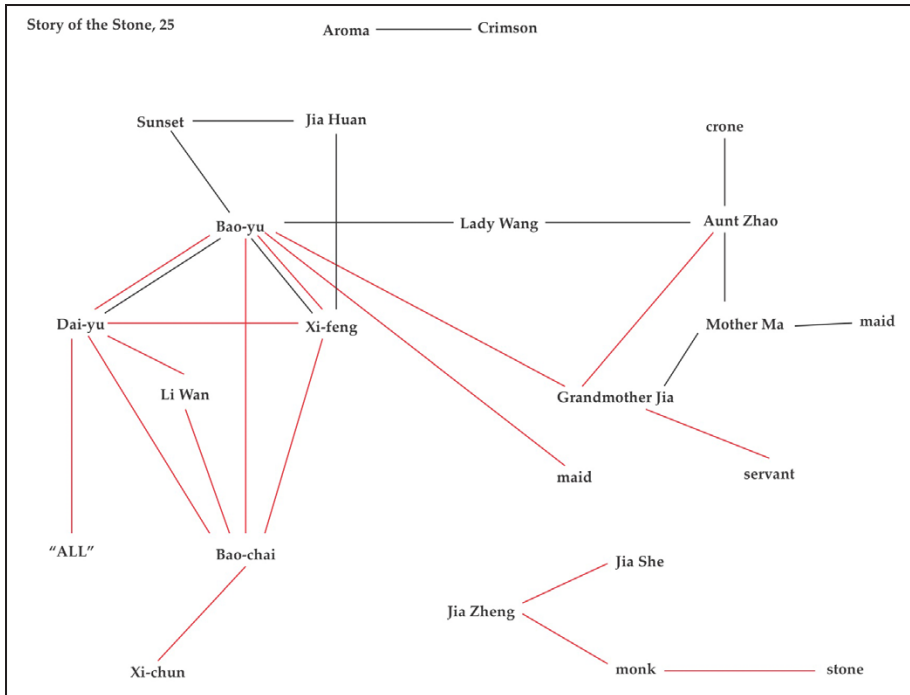


Figure 51

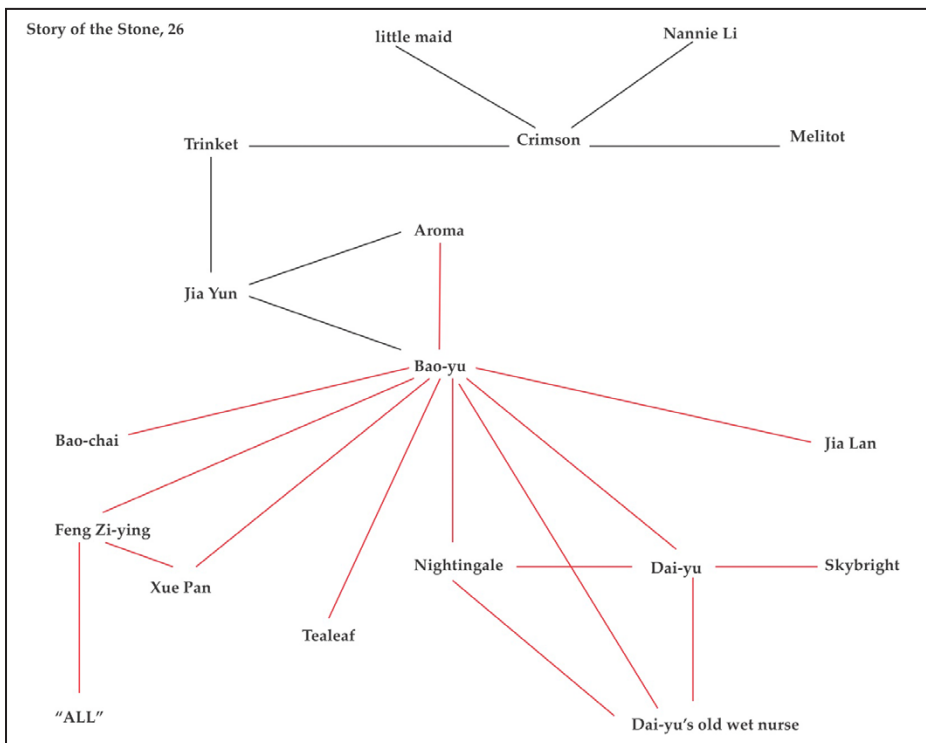


Figure 52

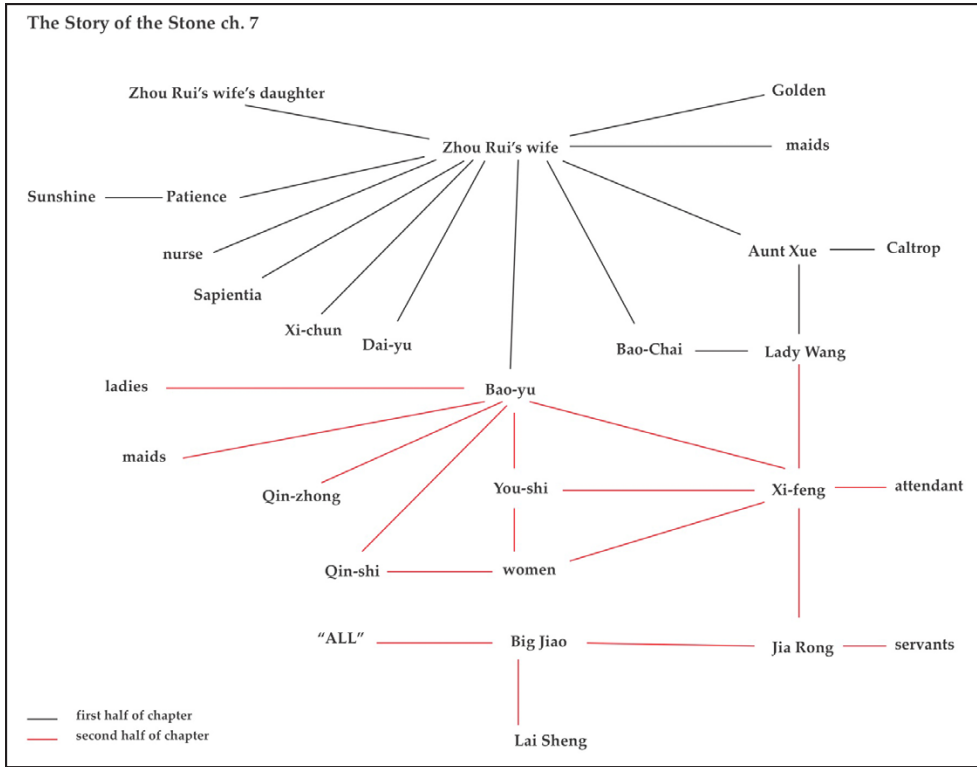


Figure 53

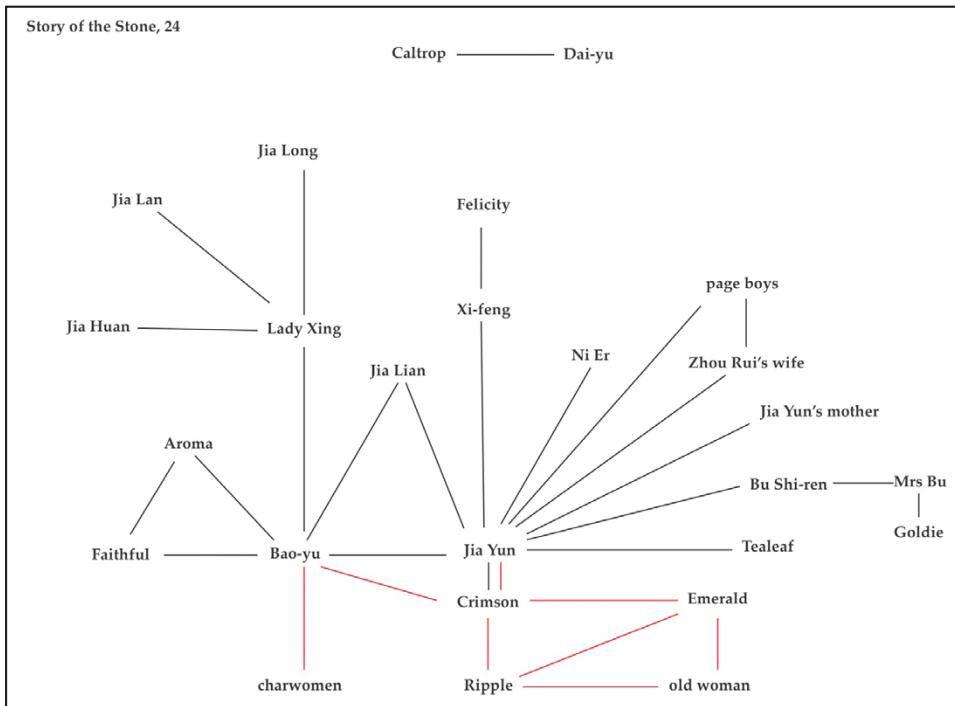


Figure 54

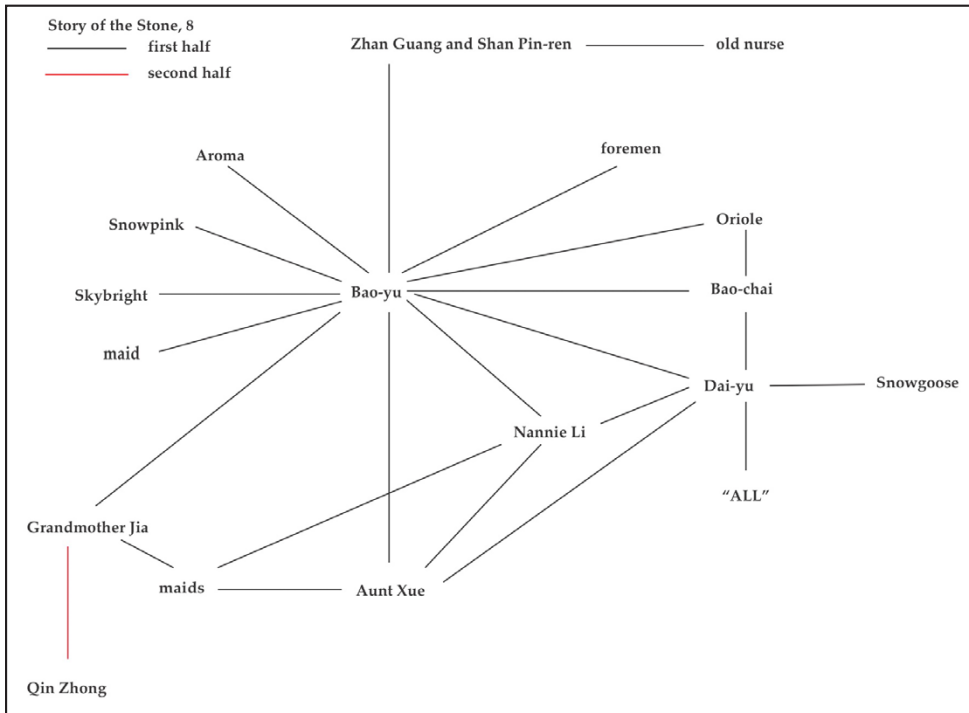


Figure 55

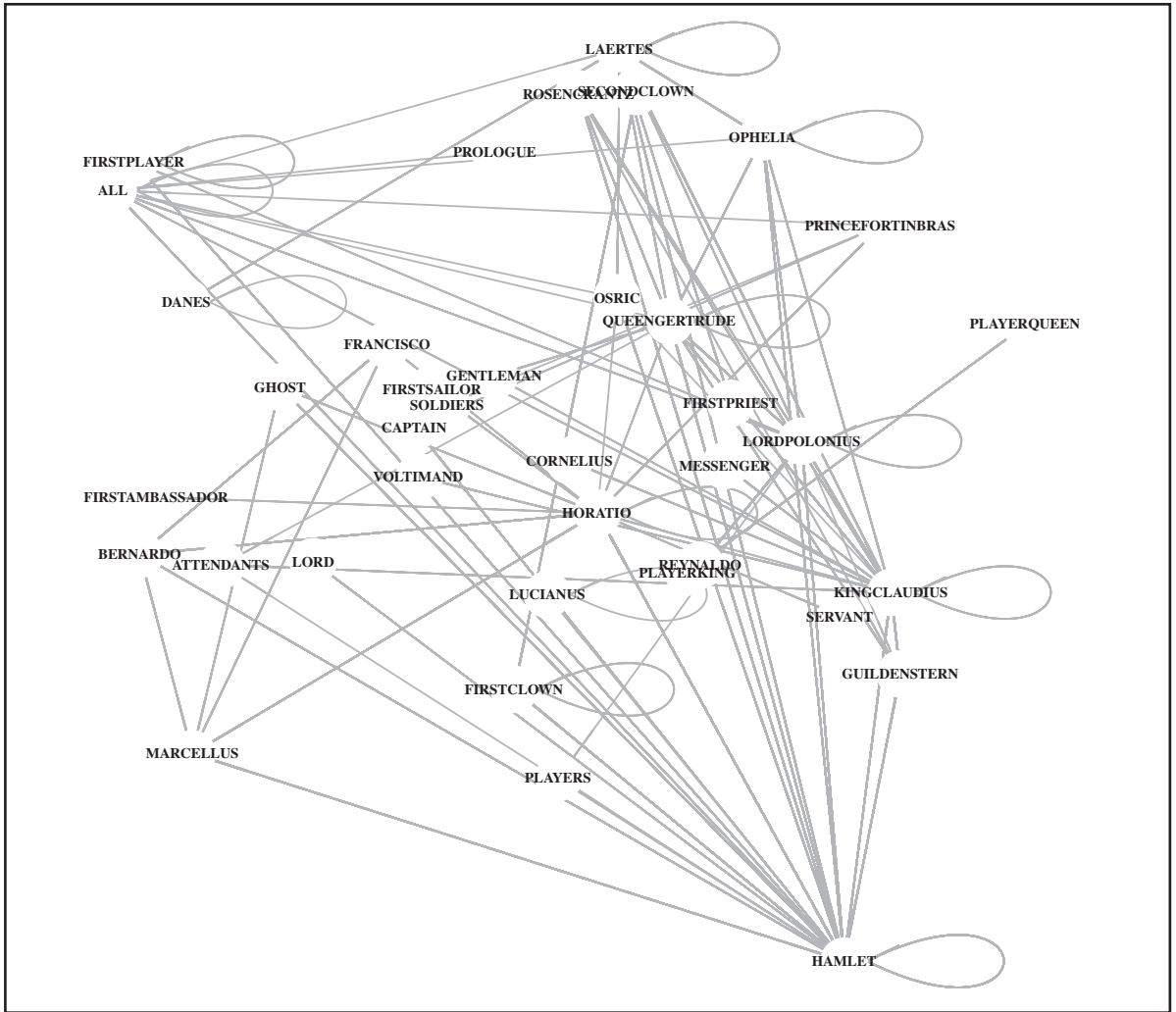


Figure 56

