Network Analysis of Dramatic Texts

Session 2: Dialogue Networks, Directed Graphs, and Comparative Analysis

Introduction

Last session, you built a **co-presence network**—who shares scenes with whom. Today you'll build a **dialogue network**—who actually speaks to whom—and make it **directed**, capturing not just connection but the flow of communication.

You'll also learn additional ways to measure importance in a network. The goal is to see how different modeling choices reveal different aspects of your play's structure.

Before You Begin

Make sure you have:

- Your Google Colab notebook from Session 1
- Your Gemini conversation with the play uploaded
- Your co-presence network visualization and edge list

Part 1: Building the Dialogue Network (25 minutes)

A dialogue network connects characters who speak directly to each other. Unlike co-presence, this captures *active interaction*—and by making it directed, we can see who initiates communication and who receives it.

Step 1: Extract Dialogue Data

In your Gemini conversation with the play, use this prompt:

PROMPT 1: Dialogue edge list

> Now I want to build a dialogue network for this play. Two characters are connected if one directly addresses the other in speech.
>
> Generate an edge list with columns: Source, Target, Weight
>
> **Make this a DIRECTED network:**
>
> - Source is the speaker
>
> - Target is the person being addressed
>
> - Weight is the number of separate speeches in which Source addresses Target
>
> Only include direct address (speaking TO someone), not speaking ABOUT someone who isn't present.

Step 2: Spot-Check

This extraction is harder than co-presence, and Gemini *will* make mistakes. Verify at least one scene in detail:

PROMPT 1b: Verification

> In [Act X, Scene Y], walk me through who speaks to whom. For each speech, identify: (1) who is speaking, (2) who they are addressing. Note any speeches where the addressee is ambiguous.

Compare this to the text. Common errors to watch for:

- Soliloquies counted as dialogue (they shouldn't be—no addressee)
- Speeches to groups attributed to one person
- Asides or speeches to the audience counted as dialogue

Record Your Verification

What scene did you check? What errors (if any) did you find and correct?

What About Characters Who Are Mentioned But Never Appear?

A third kind of network: the mention network

Neither co-presence nor dialogue networks capture characters who are talked *about* but never appear—or rarely appear. Think of Godot, who structures the entire play without ever arriving. Or dead characters in revenge tragedies. Or offstage powers (kings, gods, absent parents) who shape action through their influence rather than their presence.

We won't build a mention network today, but keep this in mind: when characters discuss someone who isn't there, that's dramaturgically significant information that your network doesn't capture. In your analysis, note who is frequently mentioned but absent or peripheral.

Part 2: Visualizing Directed Networks (20 minutes)

Now you'll visualize your dialogue network. Because it's **directed**, the visualization will show arrows indicating who speaks to whom.

PROMPT 2: Directed network visualization

Generate Python code using NetworkX and Matplotlib to visualize my dialogue network as a DIRECTED graph. The code should:

1. Read the directed edge list CSV (I'll paste it as a string)
2. Create a directed graph (DiGraph)
3. Use arrows to show direction of speech
4. Size nodes by total degree (in-degree + out-degree)
5. Make edge thickness proportional to weight
6. Use a spring layout
7. Label all nodes
8. Print a table showing: character name, in-degree, out-degree, total degree—sorted by total degree descending

Make the figure at least 14x14 inches so arrows are visible.

Run this in Google Colab. You should see arrows between characters, with thickness indicating how often one addresses the other.

Reading the Visualization

Look for:

- **Asymmetrical relationships:** thick arrow one way, thin or none the other
- **High out-degree characters:** many arrows pointing away (they address many people)
- **High in-degree characters:** many arrows pointing toward them (they are addressed by many)
- **Imbalanced characters:** big difference between in-degree and out-degree

Part 3: Calculating Multiple Centrality Measures (15 minutes)

Degree centrality tells you how *many* connections a character has. But there are other ways to measure importance:

Measure	What it captures
Degree	How many connections. Breadth of involvement.
Betweenness	How often you lie on the shortest path between others. Bridging, brokering, control of information flow.
Closeness	How quickly you can reach everyone. Access, reach, being well-positioned.
Eigenvector	Are you connected to other well-connected people? Proximity to power.

PROMPT 3: Centrality comparison table

> Add to my Python code: calculate and display a table comparing these centrality measures for each character in the dialogue network:
>
> - Degree centrality (using the undirected version of the graph)
> - In-degree and out-degree separately
> - Betweenness centrality
> - Eigenvector centrality (use the undirected version)
>
> Sort by degree centrality. Show the top 12 characters. Round values to 3 decimal places.

Analyzing Centrality Divergences

Look at your centrality table and answer:

1. Which characters rank high on ALL measures? (These are unambiguously central.)

2. Find a character with high betweenness but lower degree. What does this suggest about their structural role?

3. Find a character with high eigenvector but lower degree. What does this suggest?

4. Compare in-degree vs. out-degree for 2-3 characters. Who speaks more than they're spoken to? Who is spoken to more than they speak? What might this reveal about power, status, or desire?

Part 4: Comparing Co-presence and Dialogue Networks (15 minutes)

Now for the key analytical move: comparing the two networks you've built. The *difference* between them often reveals the most interesting patterns.

PROMPT 4: Network comparison

> Generate Python code to compare my co-presence network and dialogue network. The code should:
>
> 1. Load both edge lists
> 2. Create a comparison table showing each character's degree centrality rank in BOTH networks
> 3. Calculate the rank difference (co-presence rank minus dialogue rank)
> 4. Identify: (a) characters who are MORE central in co-presence than dialogue, (b) characters who are MORE central in dialogue than co-presence
> 5. Create a side-by-side visualization showing both networks with the same node positions

Interpreting the Comparison

1. Who is more central in co-presence than dialogue? (Present often, but speaks little or to few people)
What might this reveal about their role (servants? observers? silent witnesses?):

2. Who is more central in dialogue than co-presence? (Speaks to many, but in fewer scenes)
What might this reveal (concentrated influence? intense scenes?):

3. Are there edges in co-presence that don't exist in dialogue? (Characters who share scenes but never speak to each other)
Why might the playwright have them present but silent:

4. Who is frequently mentioned but absent or peripheral in your networks?
Think about characters discussed by others who don't appear (or rarely appear). What influence do they exert despite their absence?

Part 5: Choosing Your Focus for the Write-Up (10 minutes)

For your write-up, you need a specific analytical focus—not just description of the network, but an *interpretive argument* that uses network analysis as evidence.

Possible Directions

- **A bridge character:** Someone with high betweenness who connects otherwise separate groups. How does their structural position relate to their role in the plot?
- **An asymmetrical relationship:** Two characters where one speaks to the other far more than the reverse. What does this reveal about power, desire, or status?
- **The gap between presence and speech:** A character whose centrality differs dramatically between co-presence and dialogue networks. What dramaturgical purpose does this serve?

- **Absent influence:** A character who is frequently mentioned but rarely or never present. How do they shape the action without appearing?
- **Network structure and genre:** How does your play's network structure reflect its genre (tragedy vs. comedy, ensemble vs. protagonist-driven)?

Draft Your Research Question

Write one interpretive question your network analysis can help answer:

Before You Leave

1. Save your updated Colab notebook with all visualizations
2. Screenshot or export your key visualizations (both networks, centrality table)
3. Save your dialogue edge list
4. Note your research question for the write-up

Write-Up Guidelines (Due: 1 hour outside class)

Structure your write-up as follows:

1. Modeling Choices (1 paragraph)
Briefly explain the two networks you built (co-presence and dialogue). What does each capture? What does each miss? Why does this matter for interpretation?

2. Key Findings (2 paragraphs)
Present 2-3 specific findings from your centrality analysis and network comparison. For each:

1. State what the data shows
2. Interpret what it reveals about the play
3. Ground your interpretation in specific moments from the text

3. Limitations and Extensions (1 paragraph)
What doesn't your network capture? (Absent characters who are mentioned? Tone and content of speech? Change over the course of the play?) What would you do differently with more time?

Include:
1. At least one network visualization
2. At least one centrality table or comparison

Quick Reference: Interpreting Centrality Divergences

Pattern	Possible interpretation
High betweenness, low degree	Bridge figure: few connections, but structurally crucial. Schemers, go-betweens, messengers.
High eigenvector, low degree	Proximity to power: few connections, but connected to important people. Confidants, advisors.
High out-degree, low in-degree	Addresses many, addressed by few. Could indicate unreciprocated desire, lower status, or isolation.

High in-degree, low out-degree	Addressed by many, addresses few. High status, authority figures, objects of attention.
Higher in co-presence than dialogue	Present often, speaks little. Silent witnesses, servants, observers, background figures.
Higher in dialogue than co-presence	Active in fewer scenes, but dominates when present. Concentrated intensity.

The key insight: Different measures operationalize different concepts of importance. A character can be central in one sense and peripheral in another—and that divergence is often where the most interesting interpretation happens.