Socio-Semantic Network Motifs Framework for Discourse Analysis

Bodong Chen, Xinran Zhu, Hong Shui
University of Minnesota

Collaborative Discourse?

- **Socio-cognitive learning theories**: *Interpersonal communication* is essential for learning in various contexts.
- In the context of education: Collaborative discourse aims to leverage both cognitive and social process for learners to engage in activities such as articulation, explanation, questioning and knowledge co-construction.

Collaborative Discourse

Learners' interpersonal communication and intersubjective meaning-making to achieve learning goals beyond each individual.

A direct exchange between humans who can contribute intentionality and understanding to one another – the foundational act of language & an important tool for learning.

In authentic settings, sophisticated collaborative discourse can involve complex dynamics of social and cognitive processes.

Analysis of collaborative discourse - a multi-faceted phenomenon

Analytic approaches in three domains:

Cognitive Domain
– specific constructs of cognition

Content analysis

Social Domain
- constructs related to group dynamics, coordination, and affective factors

Content analysis & Social Network Analysis (SNA)

Integrated Domain
- connection between cognitive and social domains, e.g., transactivity.

Network approaches:

Content entities such as words can be studied as networks (e.g., co-occurrence, word sequence).

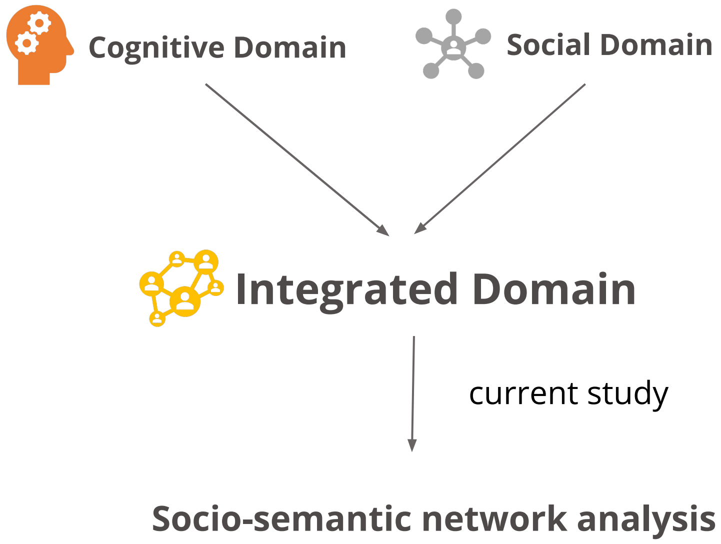
Social network centrality measures are used to evaluate a student's social position in a class & network-level measures.

Socio-semantic network analysis integrates semantic features of discourse with social networks.

(Chen, Andrews, Hmelo-Silver, & D'Angelo, 2019)

Challenges of the network approaches

- Researchers need to more explicitly address discourse processes and assumptions when constructing network from discourse data
- Given the close-knit relationship between cognitive and social aspects of discourse, we need more ways to examine the integrated domain of discourse
- We need to develop more actionable discourse indicators to improve the impact of discourse analytics



(Chen & Poquet, 2022)

The current study

Cognitive Domain

Social Domain

Integrated Domain

Current Study

Socio-semantic network (SSN) motifs framework

Grounded in socio-cognitive learning theories

Inspired by advances in network science

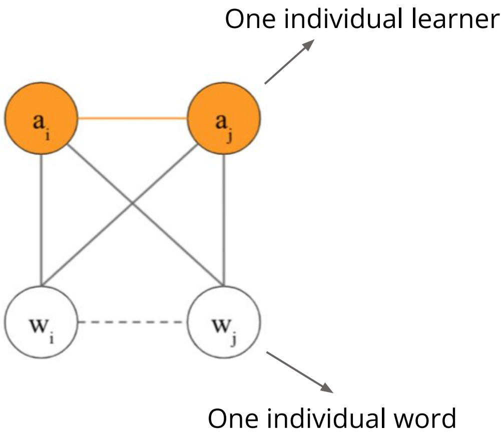
Motivated by a need for more integrated approaches to investigating learning dialogues

The Socio-Semantic Network Motifs Framework: Overview

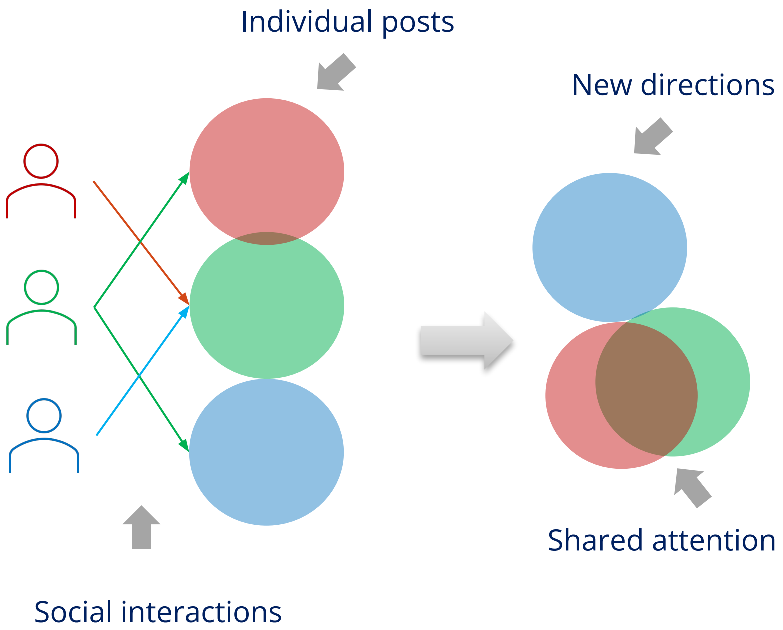
- Discourse as social-semantic networks (SSNs)

- The SSN motifs: minimal sets of social and semantic entities that are basic building blocks of a socio-semantic network.

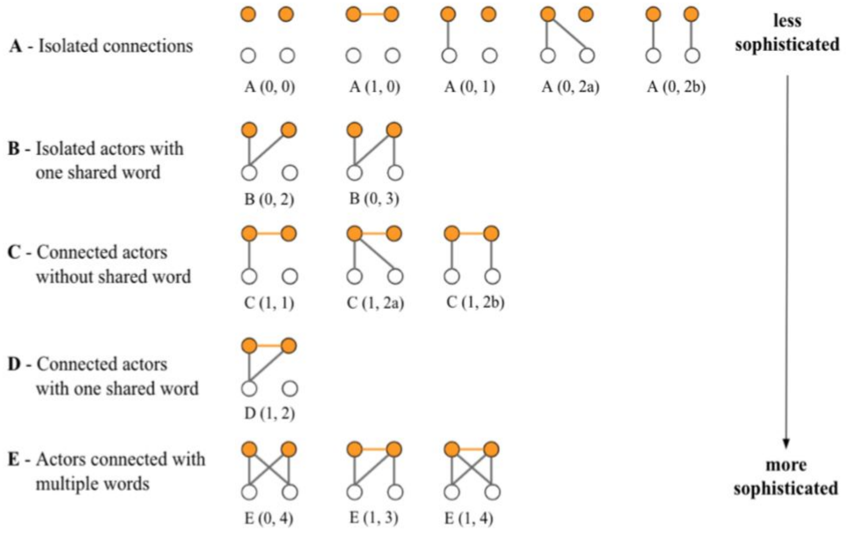
- In network science, network motifs have been widely used to examine a variety of networks including biological, technological, infrastructural, and social networks (e.g. the two-layer network motifs for the socio-ecological systems).



The SSN Motifs: Situating in Collaborative Discourse

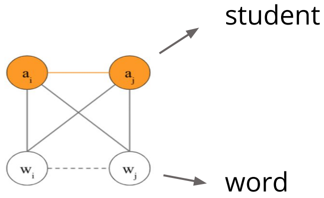


Modeling Collaborative Discourse



Modeling Collaborative Discourse as SSN – a case study

- Analytical decisions - what information to be retained and discarded?
    - the top 100 high frequency words that have appeared for minimally 5 times are incorporated in the socio-semantic network.
    - Define the nodes and edges:



Word-Word
- Not considered in this study

Learner-Word
- Only writing behaviors were considered
- Threshold: edges with a weight 2.

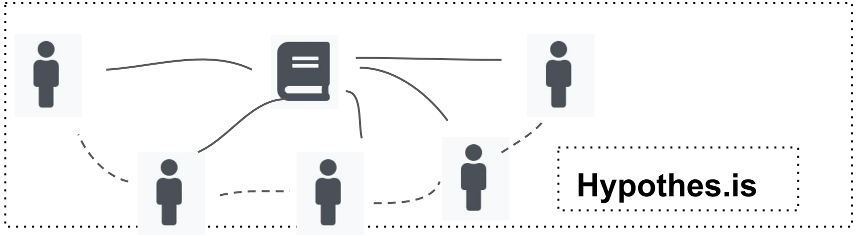
Learner-Learner
- Undirected for simplicity

Computing SSN Motifs and the Significance Profile

- Whether the generated SSN is significantly different from random graphs?
  - Step 1: compute motif frequencies (*motifr* R package, v 0.5.0) – initial indicators of the discourse's motif profile
  - Step 2: examine the significance of the motif frequencies
    - Generate 1,000 refined Erdos-Rényi random graphs as the null model
    - Compare the empirical network's motif frequencies with the random graphs
    - A *Z*-score (-1, 1) is calculated for each SSN motif to show its over- or under-representation in the empirical network

Application of the SSN Motifs Framework – A Study of Social Annotation

- **Context**: a secondary dataset generated from an undergraduate online course



- Students (n = 13)
- Annotate 1-2 readings each week and reply to each other.
- 18 readings across 11 weeks

Quantity of posts

- In total:
  - 478 Hypothes.is annotations
  - 469 replies
- On average:
  - Each reading had 26.6 annotations (SD = 2.6) and 26.1 replies (SD = 4.0).

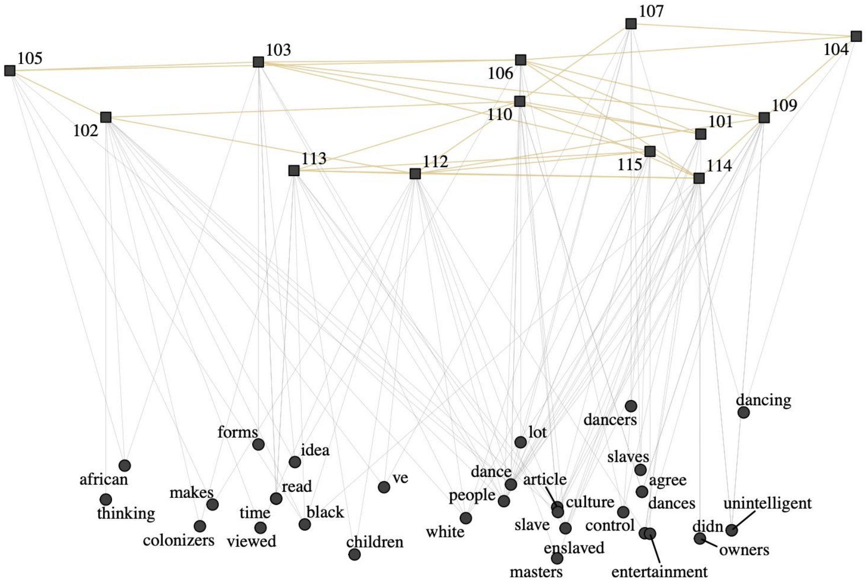
Quality of posts (human-coding)

A four-level knowledge construction coding-scheme comprising (1) Initiation, (2) Exploration, (3) Negotiation, (4) Co-construction (Zhu, et al., 2021).

- higher levels → higher order-thinking skills
- 2.36 (SD = 0.21)

Application of the SSN Motifs Framework – A Study of Social Annotation

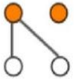
- **Findings** – An example socio-semantic network created from discourse around a particular reading.



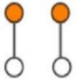
- **Upper layer**: the undirected interaction network of students
- **Lower layer**: high frequency words generated from students' written discourse in a given week
- **Link between two layers**: a word was mentioned at least twice in a student's posts.

Application of the SSN Motifs Framework – A Study of Social Annotation

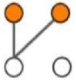
- **Findings** – Motif Analysis
    - Overview of the significance of SSN motifs across all readings



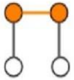
A (0, 2a)



A (0, 2b)



B (0, 2)



C (1, 2b)

E (1, 4)

Motif	A(0,0)	A(1,0)	A(0,1)	A(0,2a)	A(0,2b)	B(0,2)	B(0,3)	C(1,1)	C(1,2a)	C(1,2b)	D(1,2)	E(0,4)	E(1,3)	E(1,4)
Mean	0.22	-0.04	0.01	0.33	-0.49	-0.21	-0.07	-0.13	-0.02	-0.21	0.12	0.08	0.18	0.39
SD	0.08	0.15	0.06	0.25	0.20	0.26	0.07	0.03	0.17	0.15	0.15	0.10	0.07	0.09

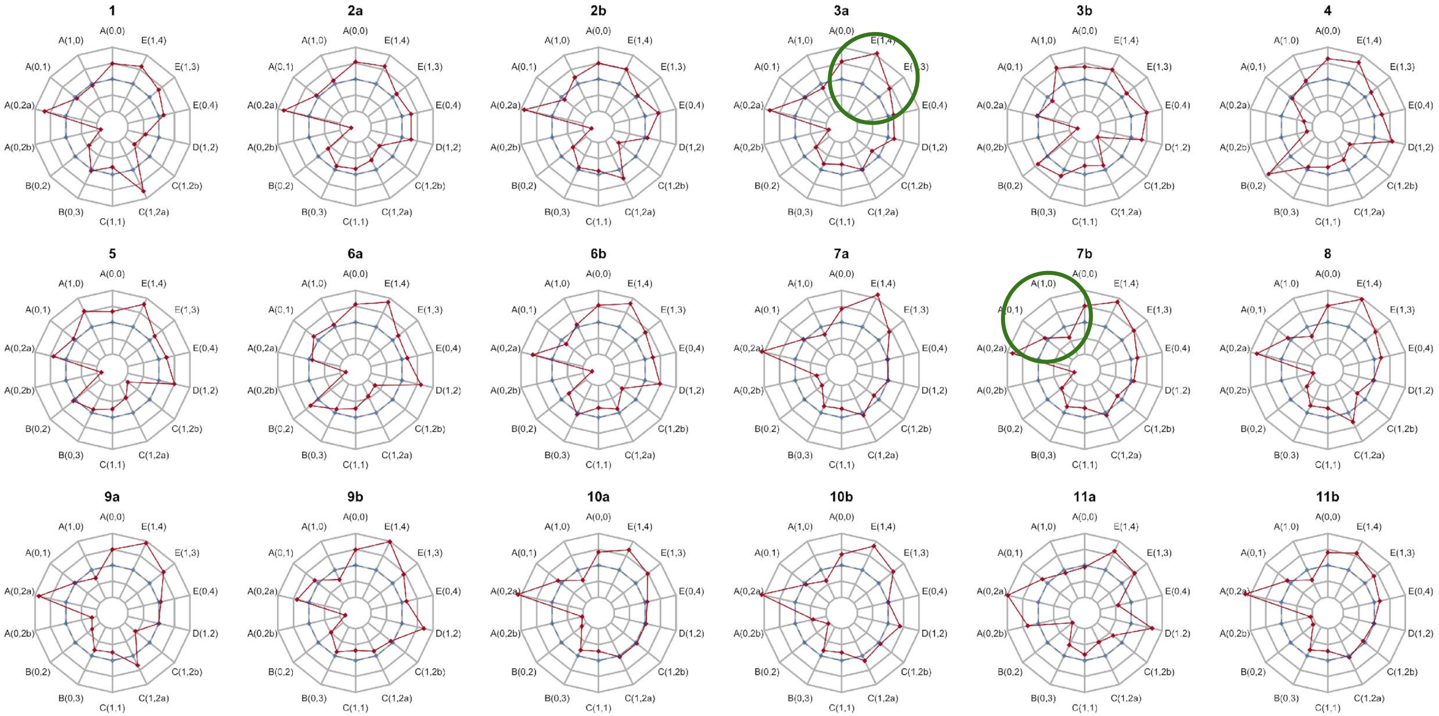
Over-represented

Under-represented

    - Interpretation: it is insufficient to evaluate discourse based on one SSN motif; instead, it is critical to concurrently consider multiple motifs.

Application of the SSN Motifs Framework – A Study of Social Annotation

- **Findings** – visualization of the SSN Motif Profiles of readings



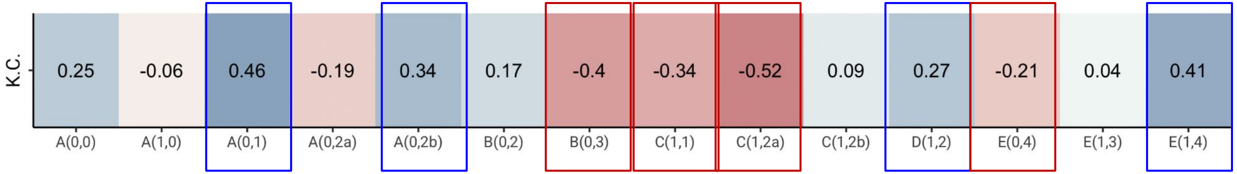
Two ways to look at the graphs:

1. In general, the class was well-connected: over-presentation of more sophisticated motifs and under-representation of less sophisticated motifs
2. Zooming into a particular reading to compare it with other readings

Application of the SSN Motifs Framework – A Study of Social Annotation

- **Findings**

  - **Correlation** (Spearman's ρ) between SSN motifs and knowledge construction:



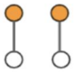
- **positively** correlated: **A(0,1), E(1,4)** (Spearman's $\rho > .4$); **A(0.2b), D(1,2)** (Spearman's $\rho > .2$)
- **negatively** correlated (Spearman's $\rho < -.2$): **B(0,3), C(1,1),** and **C(1,2a),** and **E(0,4)**

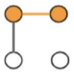
  - **Interpretation:**

    - The average normalized Z-scores of motifs A(0,1), B(0,3), C(1,2a), and E(0,4) close to zero, meaning they were nonsignificant in comparison with the random graph baseline.

Motif	A(0,0)	A(1,0)	A(0,1)	A(0,2a)	A(0,2b)	B(0,2)	B(0,3)	C(1,1)	C(1,2a)	C(1,2b)	D(1,2)	E(0,4)	E(1,3)	E(1,4)
Mean	0.22	-0.04	0.01	0.33	-0.49	-0.21	-0.07	-0.13	-0.02	-0.21	0.12	0.08	0.18	0.39
SD	0.08	0.15	0.06	0.25	0.20	0.26	0.07	0.03	0.17	0.15	0.15	0.10	0.07	0.09

    - Therefore, higher Z-scores of **E(1,4), and A(0,2b)** were associated with higher knowledge construction, whereas **higher C(1,1)** were linked to lower knowledge constructions.


A (0, 2b)


C (1, 1)


E (1, 4)

Conclusion and Implications

- Key findings:
  - We proposes a nascent socio-semantic network (SSN) motifs framework for the analysis of collaborative discourse.
  - Results showed general characteristics of discourse in the class as well as distinct motif profiles of different discourse segments.
  - Some SSN motifs were associated with higher- or lower-level knowledge construction

- Implications:
  - The motifs framework can provide an overview of discourse
  - In comparison with traditional descriptive statistics and SNA metrics, the SSN motifs can provide nuanced information about discourse, which can be used to evaluate instruction and inform pedagogical actions.
  - The framework is generic enough to be adapted to different discourse contexts.

Future Work

- To further refine the motif classification system
- Apply the framework to other discourse contexts and situate it closely in pedagogical designs
- Combine SSN motif analysis with other analytical methods
- Explore pedagogical interventions based on network motifs, such as “critical gaps” showing high impact links that could create a large number of sophisticated motifs.

Thank you!

Learning Futures Research Group:
https://learningfutures.github.io/

@bod0ng - chenbd@umn.edu
@XinranZ1 - zhu00323@umn.edu
@shui_hong - shui0003@umn.edu

Questions and Suggestions

University of Minnesota
Driven to Discover®