Exposing Chat Features Through Analysis of Uptake Between Contributions

Daniel D. Suthers  
University of Hawai‘i  
suthers@hawaii.edu

Caterina Desiato  
University of Hawai‘i  
cdesiato@hawaii.edu

Abstract
Understanding distributed learning and knowledge creation requires multi-level analysis of local activity and of how this local activity gives rise to larger phenomena in a network. Computational support is needed for such analyses due to the size of the data and distributed nature of interaction. This paper reports on one step towards implementing an analytic framework that addresses these needs. Contingencies, defined as observed relationships between contributions that evidence contextual relevance, are computed according to automatable rules, and combined to infer uptake relations between contributions. The resulting uptake structure is then analyzed through various network-analytic methods and is also transformed into a graph of uptake between actors for social network analysis. Our initial results show that a simple contingency analysis based on temporal factors, actor addressing, and lexical overlap provides structures of sufficient quality for identification of major features of a discussion and the roles of actors. The results are expected to improve as semantic analysis is added.

1. Introduction

Administrators, facilitators and researchers of large online learning or professional networks may have questions such as: Where are the most engaged discussions taking place? Who are the central actors in these discussions, not only in terms of number of contributions but also by prompting and integrating others’ contributions? What ideas receive the most development? A satisfying answer to these questions requires in-depth examination and interpretation of the interaction that took place. In particular, the sequential structure of the interaction [14] is relevant to these questions. Engagement is displayed when actors take up each other’s contributions. Central actors can be identified by how their contributions are taken up by others. Identification of the development of ideas requires tracing out threads of discussion. But human time and ‘processing capacity’, as it were, is limited. With so much happening, where should we focus our efforts? Is it possible to get preliminary answers to these questions through automated means? These are some of the questions we are beginning to address with an analytic approach we are developing [17].

The general idea is that we can approximate relevant sequential structure of interaction by observing ways in which contributions are contingent upon each other. Contingencies are empirical relationships between acts, such as shared media, mutual relationships in time and space, and relationships between the content produced or manipulated. Clusters of contingencies indicate that there may be a non-accidental relationship between contributions that we call uptake. Uptake is when an actor takes some aspect of a prior contribution as being relevant for that actor’s ongoing activity [16].

We are developing algorithms for automated identification of empirical relationships (contingencies) between contributions and conversion into sequential structure of uptake between contributions to narrow down the candidate answers to the above questions. These algorithms provide uptake models of discussions that can be analyzed in various ways to be discussed. Human analysis is then required for final assessment, but the amount of data to examine is greatly reduced. We have been comparing rule-based uptake analysis with human analysis as part of our formative evaluation for improving the algorithms. However, the uptake model is not the end product of interest and we are not primarily claiming that the algorithms correctly identify uptake. Rather, our interest is in providing useful information concerning the location of engaged discussions, the roles of actors, and the development of ideas; and our claim is that the uptake model identified by the algorithms is sufficient to identify good candidates for further human analyses towards these ends. We illustrate some of the applications of the rule-based uptake analysis and compare its results to human interpretation of the same discussion.

The remainder of the paper is organized as follows. After summarizing the general motivations and context of this research, we describe the virtual environment we are studying, SRI’s Tapped In environment for education professionals, and motivate this paper with questions we have about that environment. We then summarize our theoretical and analytic approaches (detailed in other publications),
and how the latter is operationalized in this study. We then provide various results of a formative study comparing human and rule-based analysis and discuss the potential for further refinement and applications.

2. **General Motivations and Context**

The goal of this research is to develop methods for understanding large online networks, including virtual organizations and online communities. We consider socio-technical network [10] to be the more general term, including loosely affiliated networks of individuals as well as organizations and communities [7]. We intend our approach to be applicable to all these forms of networks.

We are focused on understanding how digitally mediated interaction supports movements, associations and transformation of ideas and people. Thematically the project is committed to several analytic intuitions including that the sequential organization of interaction is critical to understanding what participants are accomplishing and the various roles they are playing; that defensible interpretations of human interaction require human analysis, but automated analysis can play a significant role in helping locate episodes of interaction of potential interest, find recurring patterns, tracing out long distance influences, and generally managing very large datasets; and that the relationship between micro phenomena (e.g., small group interaction in short time spans) and meso/macro phenomena (e.g., larger group, community, network, etc. at longer time spans) [9, 11] is of particular interest and insufficiently studied.

3. **Problem Area**

Motivated by our interest and prior work in educational communities [18, 21], and enabled by our personal contacts with the relevant researchers, we have chosen to study SRI's Tapped In (tappedin.org), an international (mostly US) network of educators engaged in diverse forms of informal and formal professional development and peer support [5, 15]. According to its developers, Tapped In has hosted the content and activities of more than 20,000 education professionals annually in more than 800 user-created spaces that contain threaded discussions, shared files and URLs, text chats, an event calendar, and other tools to support collaborative work. More than 50 “tenant” organizations, including education agencies and institutions of higher education, have used Tapped In for online courses, workshops, seminars, mentoring programs, and other collaborative activities. Also, approximately 40-60 community-wide activities per month were designed by volunteer participants. Extensive data collection capabilities underlying the system captured the activity of members. SRI colleagues provided eight years of data to us.

The present paper focuses on one of the quasi-synchronous chat sessions that are so central to Tapped In. Sessions on various topics are organized on a recurring basis, e.g., monthly, as well as one-time events. Some are open public events, such as the “After School Online” sessions, and others are private to tenant organizations. There are thousands of chat sessions in our data. This raises the most fundamental question: where are the most engaged or productive chat sessions taking place? Is it possible to characterize the “signature” of an engaged chat session in such a manner that good candidates can be found elsewhere by automated analysis? The research reported in this paper is preparatory: we examine one productive session in order to improve our methods for automatic identification of such “signatures.” Other questions pertain to individual sessions, but answers can be aggregated across discussions and media. For example, who are the central actors in these discussions, not only in terms of number of contributions but also by prompting and integrating others' contributions? And what ideas receive the most development in a given session? To answer questions such as these, we need to examine the sequential organization of contributions. The next section summarizes the theoretical approach underlying our analytic framework.

4. **Theoretical Approach**

If an individual’s learning is to benefit from the presence of others, then there must be some association between the others and the learner. We might say there must be “interaction,” but common usage of the term connotes reciprocity of communication, while learning may also be enhanced by the presence of others in more indirect ways, as in “networked learning” [3, 6] “networked individualism” [2, 20] and “stigmergy” [13]. In such situations, a learner may access information or other ideational resources provided by others without there necessarily being interaction in the conventional sense, or even an awareness of the identity of the other actors.

We use the term “uptake” to refer to all the ways in which an actor takes some aspect of the contribution or trace of activity of a prior actor as being relevant for the present actor’s ongoing activity [16]. Other motivations for this term include to generalize our conception of interaction (and other forms of association) and scale up the analysis across media types, freeing it from media-specific and interaction-type specific assumptions implied by terms such as
“adjacency pair” and “reply”. The inclusiveness and minimalism of uptake offer a conceptual tool for micro level analysis of different macro phenomena, such as roles and idea development in a socio-technical network. For purpose of the present paper, uptake is relevant because, as outlined in the opening of the paper, many of the questions that interest us require examination of the sequential structure of participants’ activities in relation to each other. Uptake is evidenced by observable relationships between contributions, as discussed further in section 5.2.

5. Formative Study

In order to focus our work we decided to analyze a single chat session in detail, for development and formative evaluation of the automated analysis. Two separate branches of analysis were undertaken in this study: rule-based analysis (the rules being applied by the authors in preparation for automation) and human analysis (based on human judgment). The two were compared to each other both directly and in terms of inferences from them.

5.1 Session Analyzed

We examined sessions that had been indexed by the Tapped In team under the heading of Professional Development, and selected a session that showed genuine engagement by participants in addressing professional issues of concern to them: a “Teaching Teachers” session on mentoring in the schools. (According to our SRI colleagues, participants in these sessions agreed to make their sessions public, including their Tapped In user names, and at this writing these sessions are publicly available on Tapped In, but to comply with our IRB requirements we are using pseudonyms for the participants.) The session was scheduled for one hour. Our analysis focused on a 33-minute portion starting about 22 minutes into the session, this portion being chosen due to the quality and level of engagement of the discussion.

We now familiarize the reader with the nature of these chats and provide examples for later use. Chat contributions are numbered in our examples: contribution #1 is about 10 minutes before the beginning of the session; our analysis starts at contribution #177; and the analysis ends at contribution #347 (the last comment of the session), for 171 contributions analyzed. Some participants arrive before the official start of the session. Barbara is a regular volunteer facilitator in Tapped In. She is present for many sessions and helps the session leaders. Maria is the leader for this session. Soon after the official start of the session, Barbara asks participants to introduce themselves to each other, and participants reply. Then Maria asks participants (#72) “I usually like to start off these discussions by finding out what experiences you have had with mentoring...either having a mentor or being one...”. This is followed by some sharing of experiences. About 18 minutes into the session, Maria refocuses discussion on the question of what is required to be a good mentor: (#152) “there are many ways to approach a mentoring situation...” and (#154) “…what are some qualities that mentors should have/develop?” In reply participants briefly list various attributes. About 22 minutes in, Maria asks:

177 22:40 Maria: “so how do you develop these skills?”
178 23:01 Maria: “how do you learn to be a mentor?”

We chose to begin our analysis with #177 because of the relatively engaged discussion that followed. Two important segments of the discussion begin with Maria’s prompts at #184 and #241. We reproduce some of the discussion that follows in Tables 1 and 2.

5.2 Rule-based Analysis

Contingencies were systematically installed between contributions according to the following rules. The intent was to find empirically observable relationships between contributions that can be computed automatically yet that provide potentially useful information concerning the structure of the interaction (i.e., constitute potential evidence for uptake).

5.2.1 Scope Contingencies. The first two kinds of contingencies capture the continuity of topics among actors, and were used to control the scope of applicability of other contingencies in a manner to be described below.

Recent Temporal Context (C-Recency): Generally, contributions are relevant to (may uptake elements of) recent contributions. In spoken interaction, each utterance is likely to be (expected to be) relevant to the immediately prior utterance: they are called adjacency pairs. The ordering of contributions in quasi-synchronous chat can disrupt the adjacency assumption: this phenomenon is illustrated by 241-252 above, and is well documented [8]. Thus we cannot assume that a contribution is taking up the one before it. However, in general, participants take up recent contributions. We needed a contingency that captures the ongoing flow of conversation where topics and concerns that have been recently shared with the group are available to be taken up. No fixed time window can accurately capture the scope of relevance. Yet, since this rule-based analysis was undertaken by a
Table 1. Sample chat episode, 184-209

184 23:35 Maria: are all good teachers good mentors?
185 23:38 Andrea: some people will take a while to get to that point
186 23:42 Andrea: No..not all
187 23:51 Nancy: definitely not
188 23:55 Helen: Training can help, but I think some is personality
189 24:09 Ashley: some people are excellent teachers but are horrible mentors
190 24:09 Nancy: some great teachers can not hold a decent conversation with an adult
191 24:11 Andrea: i had to co-ops who would be awful mentors
192 24:24 Helen: Nods
193 24:27 Lisa: That is an interesting question Maria. ... I would probably say 'yes' first off, and then wonder some more
194 24:42 Maria: it is something I have thought about often Lisa
195 24:47 Andrea: I think its alot of personality
196 25:17 Lisa: one thing a mentor has to know is how to operate with a peer, and ow to be intentional about handing over, or encouraging greater independence
197 25:18 Maria: observation has made me think that it takes an extra "special ingredient" to tip the scales
198 25:34 Nancy: I think if you have the passion for teaching you will want everyone else to feel the same
199 25:35 Andrea: agree
200 25:37 Ashley: just because someone is a great teacher doesn't mean that they have great skills that are needed to be a good mentor. some adults spend so much time talking to children that they can't talk to adults.
201 25:51 Andrea: but some have a strong passion but just keep to themselves
202 26:08 Ashley: having the passion isn't always enough, it's how you convey your passion
203 26:15 Andrea: agree Ashley
204 26:32 Andrea: cuz one teacher I worked with very strong passion..but couldn't convey it
205 26:48 Ashley: you can have a passion about fighting cancer, but if you don't relay your feelings and ideas to others it's useless. same for teaching
206 26:48 Lisa: part of being a good mentor will be being able to share your practice thinking: what are the decisions you are making and why, when you are dealing with a particular child and learning problem
207 27:22 Maria: yes! very similar to the strategy called Think Aloud that we use we kids
208 27:22 Joseph: I think the reasons for making the decisions you do is one of the most important things a mentor can pass on
209 27:36 Lisa: part of being a good mentor will involve being able to ask the questions that helps another talk about their own decision making and thinking

Table 2. Sample chat episode, 241-254

241 34:58 Maria: which is the difference between assigned/structured programs and a natural development of the mentoring relationship
242 35:07 Andrea: as a teacher you almost automatically reflect on lessons and think about how you can improve
243 35:13 Lisa: can you please elaborate Maria?
244 35:33 Patricia: i agree Andrea
245 35:39 Maria: sorry...thinking "out loud"
246 35:43 Nancy: assigned I think could be having a student teacher and naturally would be working with a team inside your grade level or even school wide
247 35:53 Lisa: please think out loud some more then!
248 36:18 Betty: I wish that were true Andrea, but sometimes there are teachers who have “done it that way for years” who fail to see the need to self-evaluate or self-reflect.
249 36:28 Andrea: oh yes..
250 36:29 Maria: but wouldn't it be interesting to measure success of “assigned” mentors against mentoring relationships that develop out of need or naturally
251 36:36 Andrea: Im sure there are
252 36:44 Joseph: That would be an interesting study
253 36:58 Nancy: yeah that would
254 37:04 Ashley: In our district we are required to be evaluated by the PDAS. part of the evaluation is self reflecting on our practices

human with limited time for processing, and to control the potential explosion of contingencies, we decided to operate within a two minute window of relevance. The size of this window is partly motivated by the pace of the discussion: Tapped In chats are not as fast paced as those for other populations and settings, so a shorter window would not include sufficient context. A contingency called C-Recency was installed between

<table>
<thead>
<tr>
<th>Time</th>
<th>Name</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>196</td>
<td>Lisa:</td>
<td>part of being a good mentor will be being able to share your practice thinking: what are the decisions you are making and why, when you are dealing with a particular child and learning problem</td>
</tr>
<tr>
<td>197</td>
<td>Maria:</td>
<td>yes! very similar to the strategy called Think Aloud that we use we kids</td>
</tr>
<tr>
<td>198</td>
<td>Joseph:</td>
<td>I think the reasons for making the decisions you do is one of the most important things a mentor can pass on</td>
</tr>
<tr>
<td>199</td>
<td>Lisa:</td>
<td>part of being a good mentor will involve being able to ask the questions that helps another talk about their own decision making and thinking</td>
</tr>
</tbody>
</table>
each contribution and all those that took place in the last 120 seconds.

We excluded those contributions that were posted at the same time, as it is not possible for one contribution to be taking up one that was simultaneously posted. For example, #190 cannot be directly contingent on #189 (the parallels between them must be accounted for by their mutual orientation to prior contributions), and #208 cannot be contingent on #207. This argument could apply further back, for example, a long contribution could not be typed in one or two seconds, suggesting that the exclusion zone be a function of the length in characters of the contribution (e.g., #200 was probably mostly typed before #199 appeared). However, for simplicity we stayed with the one-second rule in the present analysis. So, for example, contribution #200, which took place at 25:37, would be declared temporally contingent upon all contributions that took place between 23:37 through 25:36, inclusive, i.e., #185-199 inclusive.

**Same Actor (C-Actor):** This contingency is motivated by the sustained agency of an actor. A person may be thinking about a topic and possibly have an agenda to pursue it, so may bring this topic up again even if it has not been discussed for a while. Also, it is generally reasonable to expect that what a person does at any give time may be related to what they did recently. Therefore we installed contingencies from a given contribution by a given ego actor to all prior contributions of the ego actor contributed in C-Reency, prioritized in order of recency, or, if there are no prior contributions by the ego actor in C-Reency, to the last contribution by that actor. For example, contribution #200 is C-Actor contingent on #189, the last contribution by Ashley.

C-Reency and C-Actor were also used to limit the scope of search of lexical contingencies to be discussed below. One may think of C-Reency as representing the set of contributions that are relevant an available for uptake by virtue of having been recently been made public, and C-Actor as representing the contributions that are relevant to a given actor by virtue of that actor's continuity of agency (private concerns, objectives, etc.). From among this combined set of potentially relevant contributions we seek other contingencies that may provide corroborating evidence of uptake.

**5.2.2 Content Contingencies.** The remaining contingencies focus on the content of the contributions, albeit in a deliberately superficial way in this study. The first two handle direct addressing by name between actors.

**Addressing a Prior Actor (C-Addr):** If a contribution addresses a prior actor, such as #203 in the sequence 200-203 previously shown, then we assert a contingency between that contribution (e.g., #203) and the most recent contributions by the named actor that are within C-Reency, listed in order of recency (e.g., Ashley's #202 and #200). It may not be clear which prior contribution was being addressed by the present contribution, and it is possible the reference is to a contribution more than two minutes back, but the intention may not be clear to participants either.

**Replying (C-Reply):** When a person is addressed by name, an obligation may be created to reply to the contribution. Therefore, if contribution X names a person and that person has contribution Y in the next 120 seconds (that is, X is in the C-Reency of Y), we install C-Reply contingencies from Y to X. For example:

<table>
<thead>
<tr>
<th>Time</th>
<th>User</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td>309</td>
<td>Lisa</td>
<td>Sharon, was the restructuring supported by some other inputs to help deliver on the collaborative atmosphere?</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>315</td>
<td>Sharon</td>
<td>they have acquired some literacy coaches and PD training during school days</td>
</tr>
</tbody>
</table>

Contribution #315 will have a C-Reply contingency on #309. Another example is #194 and #193: #194 has both C-Actor and C-Reply contingencies on #193.

**Lexical Overlap (C-Lexical):** The last contingency used in this study is overlap between lexical items. This contingency was the most difficult to develop. It was revised in iterations and is still being revised. Our intent is to install contingencies that can be computed by current natural language processing technology. The revisions were concerning what should be computed, to help us choose an appropriate technology. We began by assuming that frequent words would be filtered and focusing on nouns and verbs, using word stems to identify lexical overlap even across morphological variations. For example, #200 uses the word “teacher,” making it lexically contingent on #198 (“teaching”), #190 (“teachers”) and #189 (“teachers”); and uses the word “mentor”, resulting in lexical contingencies to #196 (“mentor”), #191 (“mentors”), and #189 (“mentors”), all of which are in C-Reency for #200. However, we found that some instances of uptake were signaled by “echoing” of other categories of words (as in Trausan-Matu’s “voices”, [19]). For example, there is a sequence of echoing “some”: 185 (“some people”), 189 (“some people”), 190 (“some great teachers”), 200 “some adults”, 201 (“but some”). This observation led
us to include other word categories. Also, sequences of words (bigrams and trigrams) should be given more weight than overlap of single lexemes. For example, match of the exact string “matching personalities” should be given more weight than separate occurrences of “match” and “personality”.

5.3 Inferring Uptake

Once contingencies had been installed between contributions, the next step was to find combinations of contingencies that, taken together, provide evidence that uptake is likely. Contingencies are objectively observable relationships between events; but uptake crosses into the realm of intention. We cannot know for sure whether someone intended to take a prior contribution further, but we can observe that we have sufficient evidence to reasonably consider this possibility.

Rules for combining contingencies to infer uptake were developed incrementally. Results from an early attempt were compared to an initial human analysis in order to identify information that may be present in the contingencies but used suboptimally. This process of iterative refinement is ongoing, but the uptake structures used for our initial comparison and reported in this paper were generated according to the following rules. In general, evidence must include at least one content contingency in the scope of a scope controlling contingency, and uptake may be inferred when such a corroboration is present. When multiple uptakes were possible, we limited the results to the four strongest. Specific rules follow:

1. The most recent C-Addr is always assumed to be uptake.
2. Convergence of two or more content contingencies is always uptake
3. If there are multiple candidate acts, the top two to four are taken:  
   - more recent candidates are preferred
   - C-Lexical candidates are ranked by number of words overlapping
   - Noun and verb phrases are preferred over other categories
   - C-Reply is weak, and is overridden by others

5.4 Human Analysis of Uptake

The human analysis was undertaken to help us understand what kinds of information the rule-based analysis might be missing, as well as the potential strengths of an automated approach. “Uptake” is not a common concept, nor do we have precise definitions or coding criteria for human analysis, so it is not appropriate to think in terms of there being a single correct analysis that we are trying to approximate through rule-based or human analysis. Rather, the human analysis provides rich interpretations of the various ways in which people take prior contributions as being relevant for ongoing activity, suggesting ways to improve the rule-based analysis. Thus, the human analysis provides a reference for comparison, but does not represent a single fixed “gold standard.” The second author conducted the human analyses. Differences were discussed with partial analyses by both the first author and Taunalei Wolfgramm, leading to some revisions. The second author’s final version was used for the human analysis.

6. Results and Discussion

The two uptake structures between contributions generated by rule-based and human analysis are shown in Figures 1 and 2, drawn with UCINet’s NetDraw [1]. We evaluate the results in two ways: first, by direct comparison, and second, by comparison of the conclusions drawn from these structures. As we shall see, the analyses differ significantly in the details but provide similar pictures when used to draw aggregate conclusions.

6.1 Conversational Structures

Certain conversational structures are apparent from the visualizations. There is a large connected component in both graphs, indicating that both analyses are able to identify ways in which the discussion is coherent. Some structures are in common. For example, in both graphs contribution 241 leads to a chain of uptake involving 243, 245 and 247, an exchange between Lisa and Maria concerning the latter's thinking aloud. Another topic thread discernible in the transcript of 241-252 is the issue of assigned versus natural formation of relationships. The rule-based analysis clearly shows how this thread continues into subsequent discussion, with a side branch concerning doing an experiment.

The human analysis found the relevance of more of the contributions (compare the lists of orphan nodes on the right). There are small disconnected components in both graphs. In the human analysis, two small graphs comprise the farewell exchanges at the end of the session. In the rule-based analysis, there are four disconnected fragments, including a portion of the farewell exchange as well as three others, reflecting the fact that rule-based analysis finds fewer uptakes. Given that the human analyst applied world knowledge while the rule-based analysis was based on simple lexical overlap and other contingencies, it is not surprising
that the two structures differ. Our objective should not be making the rule-based analysis replicate the human one, but to make the rule-based analysis capable of identifying interactional structures important for learning analytics. Yet we provide some further comparisons to understand how to improve the contingency analysis.

6.2 Significant Participant Contributions

An uptake analysis should be able to identify significant contributions in the chat. Two approaches were taken.

6.2.1 In-Degree of Contributions. A simple measure of significance is the in-degree of each contribution:
since uptake relations are from uptaker to uptakee in our representation, the arcs point to the contribution taken up, so one would expect higher in-degree to indicate greater significance or impact of the contribution. The correlation between rule-based and human analysis on in-degree was computed using Pajek [4]. The Pearson Correlation Coefficient is 0.507, suggesting moderate overlap (25% of variance in common).

Examination of the higher degree nodes shows that typically the human analysis assigns more uptake relations, for example contribution 184 has 7 human versus 3 rule-based uptakes; 235 has 5 vs. 3; 241: 8 vs. 3; 250: 7 vs. 3; and 258: 6 vs. 3. Often the rule-based analysis finds a subset of the human analysis, with the latter applying additional semantic knowledge. For example, in the rule-based analysis contribution 184 is taken up by contributions 186, 189, and 193; while the human analysis identified these three plus 187, 188, 190 and 200, based on the analyst's recognition that these constitute answers to the question posed by 184. In some cases the human analysis identifies longer distance relations than the rule-based analysis. For example, node 241 is taken up by 243, 246, and 250 in the rule-based analysis; while the human analysis identified these three plus 254, 255, 261, and 264. Contributions 254 onwards are out of the temporal scope allowed by the rule-based analysis.

Based on these results, we are currently extending the temporal scope of the rule-based analysis to five minutes, and allowing for more than four uptake relations. We are encouraged that the rule-based analysis identifies subsets of the human analysis, and expect that the addition of semantic analytic tools and stemming algorithms will greatly increase coverage.

6.2.2 Input Domain Size. In-degree is a limited measure of the impact of a contribution because it does not take into account whether a contribution led to further discussion. The extended impact of a contribution can be estimated through its input domain, the set of contributions in the directed graph from which the given node can be reached. To make values comparable across graphs the input domain is usually normalized as a proportion of the size of the graph minus one (the contribution being considered). Using Pajek, we computed the input domain for the 171 contributions in both the rule-based and human analyses, and then computed the Pearson Correlation Coefficient between these two vectors. The correlation coefficient is 0.432, indicating that about 18-19% of the variance in common. Thus the human and rule-based analyses have some overlap, but differ. This is not surprising given that the rule-based analysis had no semantic basis beyond lexical overlap. It should be possible to improve substantially on the agreement by applying currently available semantic analysis tools [12]. Yet, other results discussed here show that the rule-based analysis has utility even though different uptakes are inferred.

In social network analysis, input domain size has been critiqued as a measure of prestige because a direct endorsement of one person to another (A endorsing C) is considered a stronger endorsement than an indirect one (A endorsing B who endorses C). However, a graph of conversational contributions is different from a social network, in part due to the temporal dimension. Since contributions cannot be made simultaneously, it is not possible for all contributions in a chat to be made in a manner that directly indicate uptake from a topic initiating contribution. Also, a chain of uptakes may indicate that participants are engaged and developing their ideas more deeply, so one could argue that uptakes several steps away from an initiating contribution are at least equally if not even more indicative of the value of a contribution than first order uptakes. But when dealing with sociograms (ties between actors), a more appropriate measure of the influence of an actor is Proximity Prestige. This weights arcs in the input domain according to their distance from the actor being assessed. This measure is used in the next section.

6.3 Actor Roles

We collapsed each of the uptake graphs into sociograms by counting the number of times each actor had uptake from each other actor. (See [17] for further discussion of mapping contingency graphs to other levels of modeling.) The resulting sociograms are shown in Figures 3 and 4. We also computed the proximity prestige for each actor, using Pajek. The Pearson correlation coefficient between these two prestige vectors (rule-based and human) is 0.894, showing very high agreement between the two methods concerning who the important actors are. Also, if actors are ranked by the in-degree of their contributions, the two analyses rank the same actors in the top five, indicating that at a coarse grain the analyses support similar conclusions.

The rankings of actors on proximity prestige are shown in Table 3. Maria is the facilitator of this session, and Lisa also plays the role of facilitator in other sessions. Therefore their high prestige ranking is not surprising. The ranking of Ashley was unexpected and informative. Her contributions such as 189, 200 and 202 generated significant discussion.
This result also illustrates two important points of this paper. First, although machine and human analyses may differ at a fine granularity, this “noise” may be attenuated at coarser granularities of analysis where results are aggregated. Second, we should be more concerned with whether an automated method of analysis produces useful results in its applications rather than whether the specific intermediate representations are correlated with human analysis.

### 6.4 Engagement

Apart from identifying important actors, we should interpret what the graph and the prestige measures tell us about the session. Most actors are connected to each other via uptake, and prestige is not weighted heavily on one person (several actors have the highest ranking scores, and they are not far separated from the next level), indicating that multiple actors are engaged in the discussion.

The connectivity of both the contribution-uptake and actor-uptake graphs, the existence of long threads (large input domains) in the contribution uptake graph, and the balanced prestige of many actors corroborate our judgment that this was an engaged session. One next step in this research is to compute these metrics for other chat sessions, including sessions we judge to be less engaged. We will compare the sessions on overall graph metrics such as density, centralization of prestige (how much variance there is between actors), and average input domain size, to see which metrics best discriminate between more engaged and less engaged sessions. We are also preparing to compare these metrics to simpler ones such as number of contributions or characters generated per person, to test our claim that examination of the sequential structure of chats provides a better guide to locating higher quality sessions.

### 7. Conclusions

We motivated this work with the need to answer several major questions about online socio-technical networks using activity logs. Administrators and researchers of such environments will want to know where the most engaged discussions are taking place. We believe that the automated analysis has potential, but further work is needed to test this claim. We are also interested in the roles of different actors. Our analytic approach places actors in the context of their participation in the development of a discussion, enabling us to see their roles not only in terms of number of contributions but in terms of how they influence others' contributions. Such role analysis can be aggregated across many sessions to identify actors' more pervasive roles. We have not addressed idea development in this paper, but a promising first step is to follow long threads of uptake and identify the content shared across the contributions in these threads.
The approach has great potential to scale up. Manual application of the rule-based analysis is tedious, but now that we have refined the rules it can be automated. The EEC approach allows us to take activity logs as input, from which abstract transcripts are derived. These transcripts can unify records of activity in chats, discussions, and other media [17]. Our next steps in the development of this approach include addition of semantic analysis through natural language processing tools [12]; implementing and running the automated analysis on many chats, comparing the results for chats of different quality based on human judgment, and extension to other media used in Tapped In.

Acknowledgements
This work was supported by NSF Award #0943147. Views expressed herein do not necessarily represent the views of NSF. The authors thank SRI and Patti Schank for providing the data.

References