Reciprocity in Student Online Discussions
Shitian Shen, Jihie Kim, Jaebong Yoo
University of Southern California Information Sciences Institute
4676 Admiralty Way, Marina del Rey, CA, U.S.A
Email: shitians@usc.edu, jihie@isi.edu, yoojaebong@gmail.com

Abstract: Online discussion is a popular tool for information exchange in web-based
education. Analyses of how students interact or their contribution styles can help us
understand weaknesses or strengths of the participants. This paper presents a framework for
capturing information trading behaviors in Q&A discussions using a 'reciprocity' model. We
measure the reciprocity of a student based on (a) the degree of responses that he/she received
from other students in discussing his/her questions and (b) the degree of contributions in
discussing other students’ problems. We use a linear regression to model the overall
reciprocity rate over time, and correlate the regression coefficient of the reciprocity rate to the
student course grade. We found that although the overall reciprocity rate is not statistically
correlated to the grade, high performing students have larger reciprocity rates and help other
students more actively. We expect that the reciprocity rates revealed from discussion
participations can help instructors make online activities more balanced.

Introduction
Reciprocity, as a general concept for measuring the “trade” in social exchange, contains the common sense that
giving and taking illuminate each other (Gouldner, 1960). As one of the essential and necessary attributes for
the social community (Wellman, 1999), reciprocity indicates the members’ behavior in the community (Herring,
2001) and maintains the stable social system by encouraging the mutual beneficial exchanges and preventing
antisocial behavior (Alexander, 1987). Online learning networks, as an aspect of social life, also yield the
resource exchanges (Swan and Shea, 2005). Specifically, online discussion platform is a classical application of
learning networks, which supplies a platform for students to post problems they encounter during study as well
as to assist other students. Reciprocity allows the online discussion works effectively by promoting students to
give’ under certain circumstances and they can also obtain help when they have problems. In previous work,
reciprocity is treated as one of the “seven principles for good practice in undergraduate education” (Chickering,
1996). In addition, reciprocity belongs to one of the major motivating factors that promote individuals to
contribute to online communities (Wang, 2003).
Although reciprocity requires giving and taking simultaneously, the degree of giving and taking for a group
or an individual cannot be strictly identical. There is a common recognition that giving can generate more
altruistic and stable reciprocity while taking yields more selfish reciprocity. Taking into account of
characteristics of students, we classify them into three different groups: considerate, neutral and self-centered
groups. Given the assumption that human behavior varies based on selfishness, we compare reciprocity rates
among these three groups, which can provide insight on student knowledge sharing behavior (Soller, 2003). As
the degree of reciprocity in learning is hard to measure, we propose a simple but effective way to define the rate
between the degree of giving and the degree of taking using participation types without considering the
inequality of each problem or question. We split the students into four different performance groups based on
their final course grades. Such classification can facilitate an analysis of how does the reciprocity rate vary
among different performance groups, allowing us to examine the relation between the grade and reciprocity.
Furthermore, we examine the change of the reciprocity rate over time within the semester. Such reciprocity
trend, presenting changes in student participation styles, can assist the teacher in assessing students’
contributions, comparing the performance of individuals or groups (Kapur, 2008), and detecting the cause for
the change in participation styles.

Methodology
In this section, we first describe how we process discussion data. We then present a model of reciprocity in
Q&A discussions. We show how the model captures the information trading behavior.

Data collection and processing
Our work takes place in an undergraduate course in Computer Science department at University of Southern
California. Student grades and online discussion data have been collected from the same course taught by the
same teacher in eight recent semesters, from 2006 to 2010 school year. The discussion data contain total 1,663
discussion threads and 7,164 messages. Among 370 active users, we chose 204 users who have at least two
questions, so that we can capture the information trade trend over multiple questions. Given a thread, we
classify which user the thread belongs to, based on the user id of the initial message in the thread. The first
message in a thread usually presents a problem or a question that the user has (Feng et al., 2006). We can
estimate how much help a user obtains by counting how many responses he/she gets from other users. On the other hand, we can also assess how much help the user provides to others by counting his/her replies to other users’ questions.

**Reciprocity rate in Q&A discussions**

We examine asynchronous online Q&A forums where participants trade information by posting questions and sending answers to the questions. Reciprocity is a common activity in social life but the effectiveness of reciprocity in online education forums is not clearly observed (Edwards 2002; Halloran et al., 2002). One main reason is that relations between actors in online learning communities are not very rich especially when there is limited participation within a course. We use multiple (8) semesters’ data from the same course to capture general behavioral patterns.

We define **reciprocity rate** (reciprocityR) as the ratio of giving and taking acts of a user. A user’s message is classified as a giving act of the person when it responds to a question. A response to the user’s question is counted as a taking act by the user. In computing the degrees of giving and taking ratio at a given time, we accumulate all the giving and taking acts until that time. More specifically, the reciprocity of user \( i \) at time \( t \) when he/she posts a question can be modeled as follows:

\[
\text{reciprocityR}_i^t = \frac{\sum_{j} \text{giving}_{i,j}^t}{\sum_{j} \text{taking}_{i,j}^t}
\]

where \( t \) is time at which user \( i \) posts a question. The overall reciprocity rate of the user within a semester can be computed by accumulating all the giving and taking acts in the semester. We used min-max normalization to rescale posting time ranges from the different semesters to the equal range between 0 and 1. For example, if user \( i \) posts \( n \) questions, we can calculate \( n \) reciprocity rates for the user, which shows how the rate change over time as the user posts more questions. Specifically, \( k \)-th reciprocity rate will be the rate between accumulated giving and taking acts until \( k \)-th question. Once we have the reciprocity at each time point for each user \( i \), we can regress their reciprocity rates to capture the general trends on the time space as follows:

\[
\text{reciprocityR}_i^t = \alpha_i + \beta_i t_i
\]

where \( t \) is a value ranged between 0 and 1. In the fitted line, \( \beta_i \) is a slope called the regression coefficient.

Such behavior trend information can be related to other variables such as student performance.

**Results**

**Reciprocity by Behavior and Group**

Figure 1 shows the distribution of each user’s reciprocityR over time by his/her behavior style and performance group that he/she belongs. For performance-based grouping, we used normalized values of the course final grades since the raw grades across semesters cannot be directly comparable. We split students into four grade groups (high, high intermediate, low intermediate, and low grade group) using mean (0.82) and one standard deviation (0.08).

![Figure 1. Reciprocity Trends by grade and behavior.](image)
problems. A self-centered user posts more messages relevant to his/her own problems. Other users are considered neutral.

In Figure 1, each solid line represents reciprocityR of a user over time. A dotted line is called the “reciprocal line” since its reciprocityR is 1, i.e. the number of giving acts is equal to the number of taking acts. In the graphs in the top row (considerate group), we can observe that the reciprocities of the high and the high-intermediate groups increase while those of the low-intermediate and low groups decrease over time. That is, higher performers tend to help other students. Although lower performers help other students sporadically, they stop helping before the final exam, which is located close to 1 in the timeline. In the middle second row (neutral group), most lines are parallel to the reciprocity line but their reciprocities are close to zero. As shown in Table 1, those students asked fewer questions than other behavior groups and the average length of discussion threads is shorter than others. They seem to participate passively in the discussions resulting from the fact that; the number of message received from others (taking acts) is relatively smaller than the number of message sent to others (giving acts), as shown in Table 1.

Finally, in the graphs in the bottom row (self-centered group), reciprocities in all the performance groups increase because they present more giving acts for their own problems, which increases the numerator of reciprocityRs. Interestingly, the self-centered behavior group among the low performing students has the largest regression coefficient. They seem to concentrate on their own problems rather than others’.

<table>
<thead>
<tr>
<th>behavior</th>
<th>grade</th>
<th>N</th>
<th>reg. coef</th>
<th>#questions</th>
<th>#giving for my problems</th>
<th>#giving for others’ problems</th>
<th>#taking</th>
<th>#thread</th>
<th>#user</th>
</tr>
</thead>
<tbody>
<tr>
<td>considerate</td>
<td>A</td>
<td>18</td>
<td>1.18</td>
<td>5.72</td>
<td>2.94</td>
<td>12.33</td>
<td>9.33</td>
<td>4.40</td>
<td>2.81</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>34</td>
<td>1.11</td>
<td>6.76</td>
<td>3.12</td>
<td>11.32</td>
<td>10.53</td>
<td>3.37</td>
<td>2.59</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>17</td>
<td>-0.74</td>
<td>5.82</td>
<td>2.18</td>
<td>8.47</td>
<td>12.88</td>
<td>3.24</td>
<td>2.58</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>3</td>
<td>-1.14</td>
<td>3.00</td>
<td>0.00</td>
<td>5.00</td>
<td>5.00</td>
<td>2.47</td>
<td>2.22</td>
</tr>
<tr>
<td>neutral</td>
<td>A</td>
<td>9</td>
<td>0.00</td>
<td>2.89</td>
<td>0.00</td>
<td>0.00</td>
<td>3.56</td>
<td>2.56</td>
<td>2.37</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>23</td>
<td>0.22</td>
<td>3.57</td>
<td>0.74</td>
<td>0.74</td>
<td>5.43</td>
<td>3.38</td>
<td>2.66</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>16</td>
<td>0.04</td>
<td>3.44</td>
<td>0.56</td>
<td>0.56</td>
<td>5.00</td>
<td>2.75</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>5</td>
<td>0.22</td>
<td>2.40</td>
<td>0.20</td>
<td>0.20</td>
<td>4.00</td>
<td>2.90</td>
<td>2.65</td>
</tr>
<tr>
<td>self-centered</td>
<td>A</td>
<td>20</td>
<td>0.76</td>
<td>7.85</td>
<td>5.00</td>
<td>1.15</td>
<td>9.60</td>
<td>3.89</td>
<td>2.55</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>28</td>
<td>0.87</td>
<td>6.61</td>
<td>5.82</td>
<td>1.43</td>
<td>14.11</td>
<td>4.23</td>
<td>2.71</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>18</td>
<td>0.82</td>
<td>5.00</td>
<td>3.39</td>
<td>1.11</td>
<td>7.56</td>
<td>3.66</td>
<td>2.46</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>13</td>
<td>1.04</td>
<td>5.69</td>
<td>3.54</td>
<td>1.00</td>
<td>7.31</td>
<td>3.69</td>
<td>2.60</td>
</tr>
</tbody>
</table>

Note: reg.coef (regression coefficient) A (high), B(high-intermediate), C(low-intermediate), and D(low)

Correlation Analysis between Reciprocity Trend and Grade
To investigate the relationship between reciprocity trends and grades, we generated a fitted line for each user with his/her reciprocityR values. We used a linear regression analysis as shown in Figure 2. Each slope, which is a regression coefficient (βi), represents reciprocityR trend for user i. As summarized in Table 1, high and high-intermediate performing group have the largest average regression coefficient. The low-intermediate and the low group have only negative average regression coefficients. In the graphs in the third row (self-centered group), the average regression coefficients are positive because the number of giving acts for own problems increase. Most of the high and low performing groups are either considerate or self-centered while high-intermediate and low-intermediate group members are equally distributed among different behavior groups.

Table 2: Correlation between Regression coefficient and Grade

<table>
<thead>
<tr>
<th>Behavior</th>
<th>Users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>considerate users</td>
</tr>
<tr>
<td>regression coefficient</td>
<td>0.25*</td>
</tr>
<tr>
<td>N</td>
<td>72</td>
</tr>
</tbody>
</table>

Note: *p < 0.05

Table 2 summarizes the result of a correlation analysis between regression coefficients and course grades. As expected, overall, they are not statistically correlated to each other because regression rate changes of the neutral and self-centered behavior groups are similar regardless of their grade levels. However, only for the considerate behavior group, the regression coefficient is statistically and positively correlated to the course grades, r(70) = 0.25, p<0.05.

Conclusion and Future Work
We propose a simple but effective model that captures reciprocal behaviors in student online Q&A discussions. The model illustrates different information trading patterns among different grade groups. We utilize a
“reciprocity line” to analyze the reciprocity trends of individual users. We found that although reciprocity R is not statistically correlated with the course grade, high performing students tend to present higher reciprocity rates.

The participation styles may be related to the nature of the problem (Kapur 2007), such as difficulty and well structuredness. Besides, the change of reciprocity may be associated with the emotion the student (Muukkonen 2007). We plan to explore use of reciprocity in explaining various aspects of student online learning behavior.

![Figure 2. Regressing Reciprocity Trends by grade and behavior](image)

**References**


