CREU 2015-2016 Final Report: Project Title

Sharon Hsiao, Arizona State University
Fatima Naveed, Arizona State University

I) Goals and Purpose

Online programming discussion forums such as StackOverflow harbor huge amounts of valuable information because of their popularity as trouble-shooting sites for programmers and learners to get help and exchange domain knowledge. Social, gamified Question and Answers (Q&A) sites are increasing the engagement of knowledge providers and increasing the response rate, resulting in the popularity of such sites (Vasilescu, Serebrenik, Devanbu, & Filkov, 2014). Just in the span of two years since its foundation in 2008, more than one million questions and 2.5 million answers have been posted (Treude, Barzilay, & Storey, 2011) on StackOverflow. As this amount of information continues to grow, the complexity of navigation has also increased. This CREU project involved tapping into this reservoir of knowledge, which is particularly beneficial for code reviews, conceptual questions, and novices (Treude, Barzilay, & Storey, 2011) by creating an automated method to discern the quality and usefulness of content to facilitate programming learning in cyberspace.

In the context of a Q&A website, such as StackOverflow, there are three types of posts: question, accepted answer, and answers. Often times, there is more than one "correct" answer, especially for code review questions (Treude, Barzilay, & Storey, 2011). Even though quality indicators (votes, acceptance, favoriting) point to the correctness of a possible solution, readers have differing backgrounds and varying degrees of previous knowledge which influence which answer is the best for that particular individual. Therefore, these quality indicators can not be universally applied to all users. For example, an accepted answer may be too dense for a novice to grasp. Filtering and presenting learning-inductive content will also allow for the "good", connecting students reaching for help with those who can provide it and reduce the "ugly", mindlessly looking for solutions without learning (Sande, 2010).

The research questions addressed by this project included:

- *how can learning in programming discussion forums be detected?* In order to first create a system which can optimize learning, we need to define a set of metrics which can detect learning inductive content in programming discussion forums. Background reading and a literature review were conducted in order to gain insight into the current methods of detecting learning as well as discussion forum dynamics. We hypothesized that learning can be detected by analyzing discussion forum content using engineered features which capture syntactic and semantic characteristics of a post.
- *how can learning in programming discussion forums be modeled?* Next, a model was created based on the findings from the previous research question to train a system to automatically discern learning inductive content in programming discussion forums. This was accomplished through regression models.
- *how can the value of the model that measures learning be validated?* A user study was conducted to validate the results of the model.

II) Related Work

Many engineered features have been designed to detect the value of information in other contexts. For example, in one study the detection of the usefulness of Flickr comments was automated via a machine learning approach by identifying the most influential engineered features through regression modeling. As a result, the system was able to filter valuable social media comments on digital objects with up to 89% accuracy based on expert perspectives (Momeni, Tao, & Houben, 2013). This approach to discern the usefulness of information has yet to be utilized in aiding programming education in cyberspace.

Using this approach, we started conducting research on the first two questions presented the previous section. Sampling one year of forum posts relating to Java in StackOverflow, we have been experimenting with engineering features which include syntactic features (word count, line count of code, etc), semantic features (Q&A coherence, overall sentiment, and informativeness, concepts extracted from Topic Facet Modelling (TFM) (Hsiao & Awasthi, 2015) etc), and answerer characteristics (reputation, badges, etc). After building regression models based on the defined constructive learning activities, results have shown a correlation between some of these features and constructive learning activities. This learning activity was captured by evaluating the use of words in a post that involve comparing, contrasting, elaborating, and justifying, since these indicate constructive learning (Chi & Wylie, 2014). We revised these features in order to make them more relevant toward the programming domain.

III) Process

The engineered features can be further analyzed to figure out why they are conducive to learning, which may provide insight on the learning itself. The CREU project involved creating a constructive learning model to measure usefulness of online programming discussion form posts. After creating an appropriate model, one of the main aspects of CREU project was to conduct a user study to assess the value of the artificial intelligence approach which presents the most learning inductive content more accessibly.

The user study involved completing a set of tasks using a system which presents questions and answers from StackOverflow according to the learning-inductive model that was designed. Not only did the study include evaluating the accuracy of the predictive model, but also pre and post evaluation impacts and surveys to track learning. The user's interactions with the system were monitored (keywords searched, posts viewed, etc) to capture a more accurate snapshot of learning since analyzing how a user comes to a solution is more valuable indicator of learning and understanding than the solution itself (Piech, Sahami, Koller, Cooper, & Blikstein, 2012). Not only should the user's search yield relevant, learning-inductive answers, but also other relevant questions and their corresponding answers since previous research has implied that similarities between questions is much more important than similarity between the questions and answers (Jeon, Croft, & Lee, 2005). According to the ICAP (Interactive, Constructive, Active, Passive) learning activity framework, this approach can result in passive, active, or constructive learning depending on the nature of post viewed and its application towards the assigned task. The goal is to optimize constructive learning. Data from the user study will then be analyzed and used to adjust model accordingly.

IV) Results and Discussion

- The full model was able to successfully predict constructiveness at 0.001 level, adjusted-R2= 0.6514
- Built constructive word lexicon based on ICAP learning activity framework to perform logistic analysis to create model to explain 65.14% of users' engagement activities
- Accepted Answers more likely to contain constructive words
- Difficult Answers were bookmarked and favorited more when there were more constructive words
- Passive-proactive learning
- **Syntactic Features**
  - Accepted Answers have an overall higher informativeness
  - Accepted Answers consist of more complex words
  - Text and code length are significantly higher in Accepted Answer
- **Semantic Features**
  - Accepted Answers cover more topics
  - Accepted Answers are more coherent with the questions
- **Social Features**
  - Easier topics have more votes and favorites
  - Social features may reach limit in identifying useful content in difficult topics
  - Programming discussion forums attract novice programmers

V) Future Work

We are still in the process of finishing up the user study to validate the constructive learning model.

VI) Web Links

http://creu2015fnaveed.weebly.com/

VII) Presentations and Publications

- **Hsiao, I-H.** & Naveed, F. (2015) **Identifying Learning-Inductive Content in Programming Discussion Forums**, IEEE Frontiers in Education, Oct 21-24, 2015, El Paso, Texas, USA
- **Identifying Learning Inductive Content in Online Programming Discussion Forums** Fulton Undergraduate Research Initiative Symposium @ Arizona State University

References

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes. *Educational Psychologist, 49*(4), 219-243. doi: 10.1080/00461520.2014.965823

Hsiao, I.-H., & Awasthi, P. (2015). *Topic Facet Modeling: Visual Analytics for Online Discussion Forums.* Paper presented at the The 5th international Learning Analytics & Knowledge Conference, Marist College, Poughkeepsie, NY, USA.

Jeon, J., Croft, W. B., & Lee, J. H. (2005). *Finding similar questions in large question and answer archives*. Paper presented at the Proceedings of the 14th ACM international conference on Information and knowledge management, Bremen, Germany. http://dl.acm.org/citation.cfm?doid=1099554.1099572

Momeni, E., Tao, K. Haslhofer, B., & Houben, G.-J. (2013). *Identification of useful user comments in social media: a case study on flickr common.* Paper presented at the Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries, Indianapolis, Indiana, USA.

Piech, C., Sahami, M., Koller, D., Cooper, S., & Blikstein, P. (2012). Modeling how students learn to program. Paper presented at the Proceedings of the 43rd ACM technical symposium on Computer Science Education, Raleigh, North Carolina, USA.

Sande, C. v. d. (2010). *Free, open, online, mathematics help forums: the good, the bad, and the ugly*. Paper presented at the Proceedings of the 9th International Conference of the Learning Sciences - Volume 1, Chicago, Illinois.

Treude, C., Barzilay, O., & Storey, M. (2011, 21-28 May 2011). *How do programmers ask and answer questions on the web?: NIER track.* Paper presented at the Software Engineering (ICSE), 2011 33rd International Conference on.

Vasilescu, B., Serebrenik, A., Devanbu, P., & Filkov, V. (2014). *How social Q&A sites are changing knowledge sharing in open source software communities*. Paper presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work; social computing, Baltimore, Maryland, USA. http://dl.acm.org/citation.cfm?doid=2531602.2531659