José P. González-Brenes

Language Technologies Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 152313

joseg@cs.cmu.edu
www.josepablogonzalez.com

1 Research Interests

Event logs are useful resources to diagnose and track the functioning of computational systems such as spoken dialogue systems and reading tutors. How can we use event logs to model dialogue behavior? One option is to use conventional classifiers, like logistic regression or decision trees. However, conventional classifiers rely on a fixed-size feature vector as an input; hence, we have to decide *a priori* how many features we are going to include. But, how to map a dialogue, that may vary in number of turns, into a fixed-size feature vector?

In my research I have proposed methods that use event logs as the input of machine learning algorithms to various mining tasks such as forecasting user engagement and modeling how changes in the system affect user behavior. Although successful at these difficult data mining problems, each task required some particular effort that did not leverage to others. This effort is the manual process called "feature engineering", which devises suitable features for predicting dialogue behavior.

Designing good features can require considerable knowledge of the domain, familiarity with the dialogue system, and effort. For example, performing manual feature engineering in a previous classification task (González-Brenes and Mostow 2010) took approximately two months. Ideally, we would want to model dialogue behavior directly from spoken dialogue system logs directly, without having to extract features. In the rest of this section I will describe my work modeling Spoken Dialogue Systems from logs of event streams using different representations.

1.1 Feature Vectors

Let's illustrate the discussion with an example. Suppose we are classifying computer-student dialogues using the single feature "turn duration". One approach is to extract features from a window, either from the beginning or the end of the dialogue (González-Brenes and Mostow 2011, to appear). Figure 1 shows the duration of each of the turns in a dialogue (9s, 8s, 5s, 7s, and 6s respectively). This window enables two alternative approaches:

(i) Averaging the value of the features in the window – in our example, it would be a single feature with value 6.0, or

(ii) Having a feature for every turn – in our example, three features with values 5, 7 and 6. Once we transform dialogues into feature vectors, we can train conventional classifiers on them.

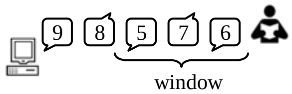

Figure 1: Dialogue described by a single feature

Approach ii may generate very large feature sets, even exceeding the number of data points. We demonstrated that a recent machine learning approach, called L_1-regularized logistic regression, is able to handle such large feature sets without over-fitting. Moreover, this rich feature set enables better classification accuracy.

1.2 Sequences

A key issue in using machine learning to classify tutorial dialogues is how to represent time-varying data. Standard classifiers input a feature vector and output its predicted label. It is possible to formulate tutorial dialogue classification problems in this way. However, a feature vector representation requires mapping a dialogue into a fixed number of features, and does not innately exploit its sequential nature. In contrast, we explored modeling dialogue as a sequence of turns, using a recent technique called Hidden Conditional Random Fields (Quattoni, Wang et al. 2007). Hence, in the example of the previous section, we would input the whole sequence (9,8,5,7,6). This procedure (González-Brenes, Duan and Mostow, in submission), allows modeling the temporal characteristics of dialogue.

1.3 Relational Learning

To bypass the labor-intensive process of feature engineering, we propose (González-Brenes, Tan and Mostow, in submission) to learn classifiers directly from a relational database of events logged by a tutor. We propose a system that searches through a space of classifiers represented as database queries, using a small set of heuristic operators.

2 Future of Spoken Dialog Research

Modeling dialogue behavior is a fundamental step towards understanding how to build better Dialogue Systems. Figure 2 describes the general process desired to model dialogue behavior: A human speaks with a dialogue agent, which logs its state and the conversation in a relational database. The relational database is transformed into a representation that allows reasoning about dialogue behavior.

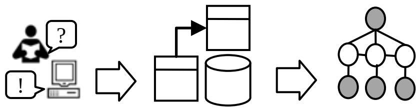

Figure 2 Predicting Behavior of Dialogue using a Relational Database

3 Suggestions for discussion

I am interested in discussing:

- How to evaluate Spoken Dialogue Systems?
- What behavior of human-computer dialogue is most interesting to model?
- What methods of Machine Learning and Human-Computer Interaction are most useful for the Spoken Dialogue community?

References

González-Brenes, J. P., A. W. Black, et al. (2009). Describing Spoken Dialogue Systems Differences. International Workshop on Spoken Dialogue Systems, Irsee, Germany.

González-Brenes, J. P. and J. Mostow (2010). Predicting Task Completion from Rich but Scarce Data. Proceedings of the 3rd International Conference on Educational Data Mining, Pittsburgh, PA.

González-Brenes, J. P., Duan, W. and J. Mostow (in submission). "How to Classify Tutorial Dialogue? Comparing Feature Vectors vs. Sequences"

González-Brenes, J. P., Tang, L. and J. Mostow (in submission). "Learning Classifiers From a Relational Database of Tutor Logs"

González-Brenes, J. P. and J. Mostow (2011, to appear). "Classifying Dialogue in High-Dimensional Space." ACM Transactions on Speech and Language Processing(Special Issue on Spoken Dialogue).

Quattoni, A., S. Wang, et al. (2007). "Hidden conditional random fields." Pattern Analysis and Machine Intelligence, IEEE Transactions on 29(10): 1848-1852.

Biographical Sketch

José currently conducts research in Project LISTEN, under the supervision of Prof. Jack Mostow. He is a PhD student in the Language Technologies Institute in Carnegie Mellon University, where he holds a Masters degree. He also holds an IMBA in Technology Management from National Tsing Hua University in Taiwan, ROC and a BSc in Computer Science from Instituto Tecnológico de Costa Rica. Besides Dialogue Systems, he is also interested in Data Mining. He recently won the RTA Data Mining challenge, against over 350 teams of all over the world, designing the most accurate algorithm to predict travel time in an Australian highway.