Melanie J. Martin   **Research Statement**

I work in the areas of artificial intelligence, information retrieval, machine learning, data mining, natural language processing, and health informatics. In addition to my own research, I work with Dr. Peter Foltz applying Latent Semantic Analysis to model discourse, and have worked on automatic recognition of subjectivity with Dr. Janyce Wiebe.

Overall Agenda

With the explosive growth of the World Wide Web has come, not just an explosion of information, but also an explosion of false, misleading and unsupported information. At the same time, the web is increasingly being used for tasks where information quality and reliability are vital, from legal and medical research by both professionals and lay people, to fact checking by journalists and research by government policy makers.

I believe we need to go beyond retrieving relevant information to be able to determine whether the information is reliable and can be verified. I define reliability as a measure of the extent to which information on a given web page can be trusted. By verifiable I mean a reasonable determination of the truth or correctness of a statement by examination, research, or comparison with similar text.

My ultimate goal is to develop a system that will automatically estimate the reliability of a given web page and allow a user to automatically verify information found on the Web.

My dissertation research focused on the reliability aspect of this problem in the medical domain. Future research will extend this work to other domains, verification of statements on a given page, and clustering of similar pages by agreement.

Reliability of Information on the World Wide Web

My dissertation research, developing a measure of the reliability of information found on medical web pages, focused on:
- Defining reliability of information in the medical domain.
- Determining features, which can be extracted from the page and the topology of the Web near the page, are good indicators of reliability.
- Learning to classify medical web pages based on their features.

To define reliability, I start with standards developed by library and information scientists (e.g. accuracy, authority, completeness, currency, objectivity), then consider document attributes specific to the Web (e.g. hub/authority, inlinks) and finally the standards from Evidence Based Medicine. (Validity of the final set of standards will be shown empirically.)

I use a variety of techniques from natural language processing, data mining and machine learning to extract features from web documents, which are indicative of reliability of medical information. My parser takes advantage of both the natural language and the html structure of the page. Some examples of features are: vector length of the page (in LSA semantic space), presence of subjective adjectives, presence of copyright, number of inlinks from outside the domain of the page.

Given the set of features, I use decision trees (C4.5), Naïve Bayes, Support Vector Machines, k-NN and clustering to learn to classify the pages. My system is able to classify medical web pages by reliability and by type of page (commercial, link, patient leaflet, etc.) with reasonable accuracy. I am currently exploring ways to improve and generalize my results.

Other Research Projects

Discourse Modeling with Latent Semantic Analysis (LSA): work with Dr. Peter Foltz (Computing Research Laboratory)

From January 2002 to present, I have worked for Dr. Peter Foltz, and as part of a research group with co-PIs, Dr. Foltz and Dr. Nancy Cooke at Arizona State East, funded by grants from ARL and ONR. My work with Dr. Foltz is on **automatic discourse analysis** in the team-communication domain. The ultimate goal is to have a real-time system that analyses the team discourse and provides feedback to team members and trainers or supervisors so that team performance can be improved on critical tasks.

My work has primarily centered on automatically discourse tagging using a predefined tag-set developed to categorized content of sequential team communication. My basic algorithm takes an utterance and uses LSA to find the most semantically similar utterances, which have previously been tagged, and estimates the most probable tag for the current utterance. I have improved the algorithm by adding some superficial syntactic features. We are in the process of investigating the correlation of tag counts and counts of tag sequences (e.g. bigrams) to overall team performance. Preliminary results confirm the findings of Bowers et al. *Results are reported in Martin and Foltz NAACL 2004. A demo of this work can be found by clicking the "Discourse Analysis" link on http://bluff.nmsu.edu/~ahmed/.*

In addition I have been involved in the hiring and supervision of other research group members. I bear primary responsibility for ensuring the integrity and consistency of the data used in LSA processing. I am currently supervising a project to annotate additional data.

Automatic Recognition of Subjectivity: work with Dr. Janyce Wiebe (University of Pittsburgh)

From May 1998 to December 2001, I worked with Dr. Janyce Wiebe and colleagues on projects aimed at **automatically recognizing subjectivity in text**. Projects included:

- Topic segmentation
- Ideological point of view
- Flame recognition

We used discourse processing techniques from computational linguistics and probabilistic classification. In probabilistic methods, we investigated probabilistic model search procedures and methods for representing lexical information as input features to machine learning algorithms.

My comprehensive exam was on the feasibility of developing an automatic system that would, given a collection of text from the Internet about a given topic, segment the text by ideological point of view.
The written portion is available at: http://www.cs.nmsu.edu/~mmartin/courses/comps_all.html

I developed annotation instructions for recognizing flames (hostile or abusive messages) in Usenet newsgroups and supervised the annotation. *Results are reported in Wiebe et al. Sigdial 2001 and Computational Linguistics 2004.* I worked with master's students to implement algorithms for automatic flame detection.