A Natural Language Processing Package for Semantic Social Referencing
Sam Spaulding
Advisor: Brian Scassellati

Abstract:
For my Senior Project I have developed a natural language processing package that analyzes text and identifies the subject of a sentence. It can recognize various word forms that refer to the same object (e.g., ‘cherry,’ ‘cherries,’ and ‘cherry pit’ are all clustered together) and computes a running ‘sentiment score’ for each detected object, based on input. I conducted the sentiment analysis in two different ways. First I tried simple lexicon-based approach. I then implemented a more sophisticated machine learning technique.

This package was designed for unfettered natural language communication, and is intended for use in a Semantic Social Referencing robotics project. For more details, see below.

Description of 1st Semester Focus
As I work on the project grew in scale and size midway through the semester I made the decision, on the advice of my advisor, to integrate this project with ROS, an open source robotics platform. As the scope of the project grew I decided to focus on the linguistic aspects of the project during the 1st semester.

Semantic Social Referencing Background

The ability to learn from a human and modify one’s own behavior is a crucial skill for artificially intelligent agents. To enable non-specialists to participate in this training process, it’s important for human-robot interactions to be as natural and intuitive as possible.

My project aims to develop a system that can reproduce one aspect of the human social learning process called “Social Referencing” on a humanoid robot. Social referencing behavior occurs when an agent learns about objects in the environment not through direct interaction, but by observing another agent interacting with the object. For instance, if a child watches an adult take a bite of food, then makes a satisfying sound and smiles, the child learns that this food is probably good to eat. If an agent observes someone reacting negatively to an object, the agent learns that the object is likewise to be rejected. This process has been simulated by an embodied robot before (see Thomaz, Berlin, and Breazeal 2005) but the verbal interaction does not make use of the semantic content of the speech. Instead, the inferred emotional content was primarily derived from prosodic information (i.e. the intonation of the speech.)

Of course, in the real world, the content of what we say provides a great deal of information about how we feel, and thus this project, "Semantic Social Referencing"
attempts to offer a different approach to the Social Referencing task by using semantic analysis tools to accomplish the task.

This project aims to develop on a humanoid robot (see Fig. 1) a system that can perform basic object recognition, map that object to a word or phrase, and then use the semantic content of a conversation to calculate a sentiment score for the object. The robot’s physical behavior will then reflect the score of the subject of discussion. This project will use the Aldebaran Nao, a small humanoid research robot, as the development system.

The three major components of this project are Object Recognition, Semantic Analysis, and Robot Behavior and Control. We need the system to recognize objects and pair them to a word, (informally, ‘carrot’ should be grounded to ‘the orange plastic thing I see in front of me’) recognize that an object is referred to in a sentence and compute a sentiment score, (‘carrots are healthy and tasty’ => ‘carrots’: ‘healthy’: ‘tasty’ => ‘positive sentiment’) and match its behavior to the sentiment score (score for ‘carrot’: +10 => ‘reach for carrot, act energetically, appear happy.’)

Users will interact with a robot in a series of ‘teaching’ and ‘discussion’ phases. In the teaching phase, the robot learns to recognize the object and builds up a sentiment for the object from the user’s verbal input. In the discussion phase, the robot will respond to prompts from the user (e.g. 'How do you feel about carrots?') with physical behavior determined by the sentiment score.

![Figure 1: Nao, a small humanoid robot](image)

**Package Description:**

The package takes input from the user until the string 'stop' is given as input. For each given sentence, the package first stems each word in the sentence. Then, each word is tagged with a part of speech. If a word is tagged as a noun and its position in the sentence indicates that it is the object of the sentence, it is added to the dictionary of ‘known objects’ if it is not already present in the dictionary.
The full sentence is then analyzed for semantic content and assigned a score, corresponding to its 'positivity.' This score is then averaged into the cumulative sentiment score for the object of the sentence.

Stemming was initially accomplished using an implementation of the popular Porter Stemming algorithm. This approach could not correctly map singular forms of many nouns to their plural forms, so the final version uses the Wordnet Lemmatizer, a more sophisticated stemmer than can reduce plural nouns to their singular forms.

Part of Speech tagging was accomplished via the open-source NLTK POS tagger. On top of this, I added a component that incorporates positional information to determine whether or not a word is the object of a sentence.

Sentiment Analysis was initially computed via a simple lexicon based approach. I looked each word up via the WordNet API, an expansive lexicon that contains an explicit sentiment score for many words. To get the sentiment score for a sentence, I summed the individual scores for the words. Due to its conceptual simplicity, this technique failed to produce good results.

As a replacement, I used Latent Semantic Mapping (LSM) a variant of the more commonly used Latent Semantic Analysis. LSM uses supervised training to construct a classifier, which then computes a positivity score on a sentence. I used the popular IMDB database (See Pang and Li, 2008) to train the map. The training samples were movie reviews from movies given scores of 9 or 10 (for examples of positive sentiment) and movies given scores of 1 or 2 (for examples of negative sentiment).

**Work in ROS**

After I had finished development of the NLP package, I started work on developing the ROS system that would run it. I have started development on a ROS node that captures video frames from the robots camera and exports the pixel maps to OpenCV for object recognition. This work will be continued next semester.

**Reflections**

The scope of the project expanded during the course of the semester, but I believe it is for the better. Making the project ROS-compatible will give it much more flexibility, and the open-source nature of ROS will widen the impact of this project. ROS also makes the complete system integration much easier. Becoming more familiar with ROS was one of my goals for this project, and even in this limited domain, I can see that it has the potential to radically expand my capabilities in robotics. The groundwork I have laid here will be an integral part of this project going forward. I am looking forward to combining natural language processing with social robot behavior in the future.