Description

Within Social Robotics there has been a great deal of research investigating non-verbal methods of communication. Computational analysis of body posture, gaze, and often direct input via a software interface are commonly leveraged to create appropriate social responses in robotic systems. Even in systems that do make use of verbal behavior, semantic information is typically discarded in favor of prosodic interpretation. I propose to engineer a system that does make use of the semantics of speech, using a

Social Referencing is the process by which an agent learns things about objects in the world by observing others interact with that object. If an agent observes someone reacting negatively to an object, the agent learns that that 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 interaction. Of course, in the real world, what we say gives a great deal of information about how we feel, and thus this project, "Semantic Social Referencing" attempts to provide a semantic solution to Social Referencing by using sentiment analysis tools to accomplish the task.

Within the field of Natural Language Processing, many groups are working on the problem of 'sentiment analysis' (given a snippet of text, classify the emotional content of that text into one of several categories.) This classification can be as simple as a binary task (e.g 'positive' or 'negative'?)) or arbitrarily complicated emotional categories ('happy' 'sad' 'upset' 'excited' etc.)

For my senior thesis, I will be working to develop a robotic system that can perform basic object recognition, map that object to a word or phrase, and then use the semantic content of a conversation about that object to calculate a sentiment score for the object. The robot's behavior will then be determined by the score of the subject of discussion. Our initial test domain is foods, and we see applications here in helping children make healthy food choices by teaching the robot. Thus, if the subject of conversation is a doughnut and the child says 'A donut is unhealthy' the sentiment score for the 'donut' object will decrease, and the robot will act appropriately, knowing that the donut is not an object to be desired.

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.
**Tech Involved:** Aldebaran Nao, NLTK package, OpenCV, Dragon Dictate Speech Recognition Software. All other code written from scratch for Python.

**Deliverables:** Code, Final Report, demo of progress to date in December.

---

**Sample Image**

---

**Future Work**

Initially I plan to develop the Sentiment Analysis classifier via either a lexicon-based approach (Using Wordnet sentiment learned from the English Language Wikipedia) or train a classifier...
using a simple machine learning algorithm (Naïve Bayes, Max Ent, or SVMs.) These simple approached rely on the so-called “Bag of Words” assumption. This assumes that documents are fully represented by a frequency count of their constituent words.

(e.g. “I am what I am” => [“I”, 2], [“am, 2], [“what”, 1])

Clearly this assumption is not true: the order of words and their context in the sentence has a great deal to tell us about the meaning of a sentence and thus represents information that is discarded by the bag of words model. For instance, “The white blood cells attacked the virus” has a very different meaning from “The virus attacked the white blood cells” yet both have the same bag-of-words representation.

Despite its limitations, the bag-of-words model performs well in many NLP tasks. However, there are other approaches that do not rely on this model. One example is the “Deep Learning” approach pioneered by Andrew Ng and colleagues at Stanford. The word of Richard Socher et al. is one example of how Deep Learning has been used to achieve state-of-the-art success at NLP tasks, including Sentiment Analysis. As an additional facet of this work, I think it would be interesting to examine classes of phrases that the Deep Learning approach can analyze but the bag-of-words approach cannot.