

Identifying Emotions, Intentions, and Attitudes in Text

Using a Game with a Purpose

Lisa Pearl & Mark Steyvers, University of California, Irvine

Introduction

Social information in text includes **emotional states**, **intentions**, and **attitudes**.

Ex: **“Come on...you have to buy this.”**
Social Information: **Intent to persuade**

Can be used for:

- understanding the intentions behind a message’s creation
- predicting how a message will be interpreted by humans reading it

Problem: Few databases exist of text annotated with human perceptions of social information, which makes it difficult to train machine learning techniques to extract this social information.

Our Proposal

Useful observation: Humans are used to transmitting social information through language.

Potential Solution: Use **human-based computation** (Kosurokoff 2001, von Ahn 2006, among others) to leverage this ability from the population, and use it to construct a reliable database of social information in text. Specifically, use a **game with a purpose (GWAP)** (von Ahn and Dabbish 2004, von Ahn 2006, von Ahn, Kedia, and Blum 2006).

Why it could work:

- Steyvers et al. 2009: **“wisdom of the crowds”** effect shown for many knowledge domains, including human memory, problem solving, and prediction
- Snow et al. 2008: a relatively small number of non-expert annotations in natural language tasks can achieve the same results as expert annotation.

Main Questions

1. Do we observe a **“wisdom of the crowds”** effect for identifying social information in text? That is, can pooling non-expert opinion yield something more reliable?
2. Can we construct a **useful database** for social information in text using a GWAP?

How useful is the database?

To demonstrate the utility of the created database for developing computational approaches to social information identification in text, we applied a Sparse Multinomial Logistic Regression (SMLR) classifier (Krishnapuram et al. 2005) to the subset of the Crowd-Labeled messages where the intended social information was perceived (624 messages).

First-pass measure: use shallow features

- unigrams, bigrams, and trigrams
- number of word types, word tokens, and sentences
- number of exclamation marks, questions marks, and punctuation marks
- average sentence and word length
- word type to word token ratio
- average word log frequency for words appearing more than once in the database

Justification: Prior research involving linguistic cues for identifying information in text has often used word-level cues (Anolli, Balconi, & Ciceri 2002, Pang, Lee, & Vaithyanathan 2002, Turney 2002, Zhou et al. 2004, Gupta & Skillicorn 2006)

The SMLR classifier model was trained to produce the label (one of eight) corresponding to the generated social information using all the features as input. Using a 10-fold cross-validation procedure, the model was trained on 90% of the messages and tested on the remaining 10%.

Confusion matrix for the SMLR classifier. The rows represent the intended social information for a message while the columns represent the labeled social information, averaged over messages and participants.

	deception	politeness	rudeness	embarrassment	confidence	disbelief	formality	persuading
deception	.36	.08	.19	.08	.08	.09	.06	.08
politeness	.05	.49	.12	.12	.05	.01	.12	.05
rudeness	.06	.06	.63	.04	.07	.07	.01	.07
embarrassment	.02	.01	.11	.76	.06	.03	.01	.00
confidence	.06	.01	.04	.08	.68	.02	.03	.08
disbelief	.08	.03	.08	.02	.09	.56	.02	.12
formality	.00	.26	.06	.03	.00	.06	.43	.15
persuading	.05	.06	.09	.03	.11	.03	.02	.61

Classifier similarities to human performance:

- **deception** (.36) and **formality** (.43) are more difficult to identify
- **confidence** (.68) and **embarrassment** (.76) are easier to identify
- **formality** often mistaken for **politeness** (.26)

Classifier mistakes:

- **deception** often mistaken for **rudeness** (.19)
- **politeness** often mistaken for **rudeness**, **embarrassment**, and **formality** (all .12)

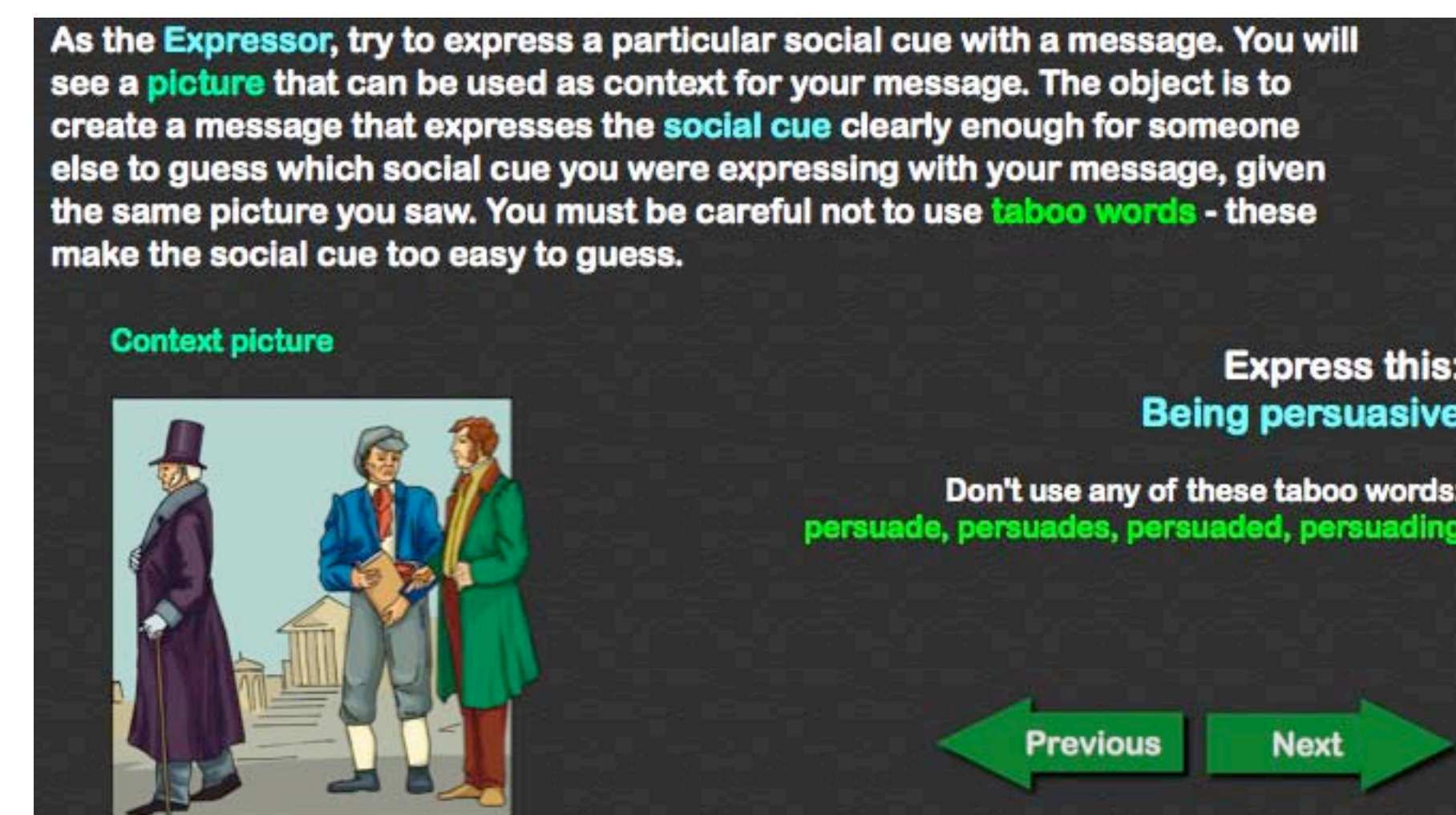
GWAP-created databases are useful

• When the classifier makes the same mistakes as humans do, this suggests humans might be using similar shallow linguistic features to make their decisions. A classifier trained on these shallow features may be able to predict which messages are likely to be confusing for humans and what linguistic features are diagnostic.

• Since the social information types we used in our GWAP can be identified automatically with some success, this suggests that these social information types are useful to pursue with automatic classification techniques.

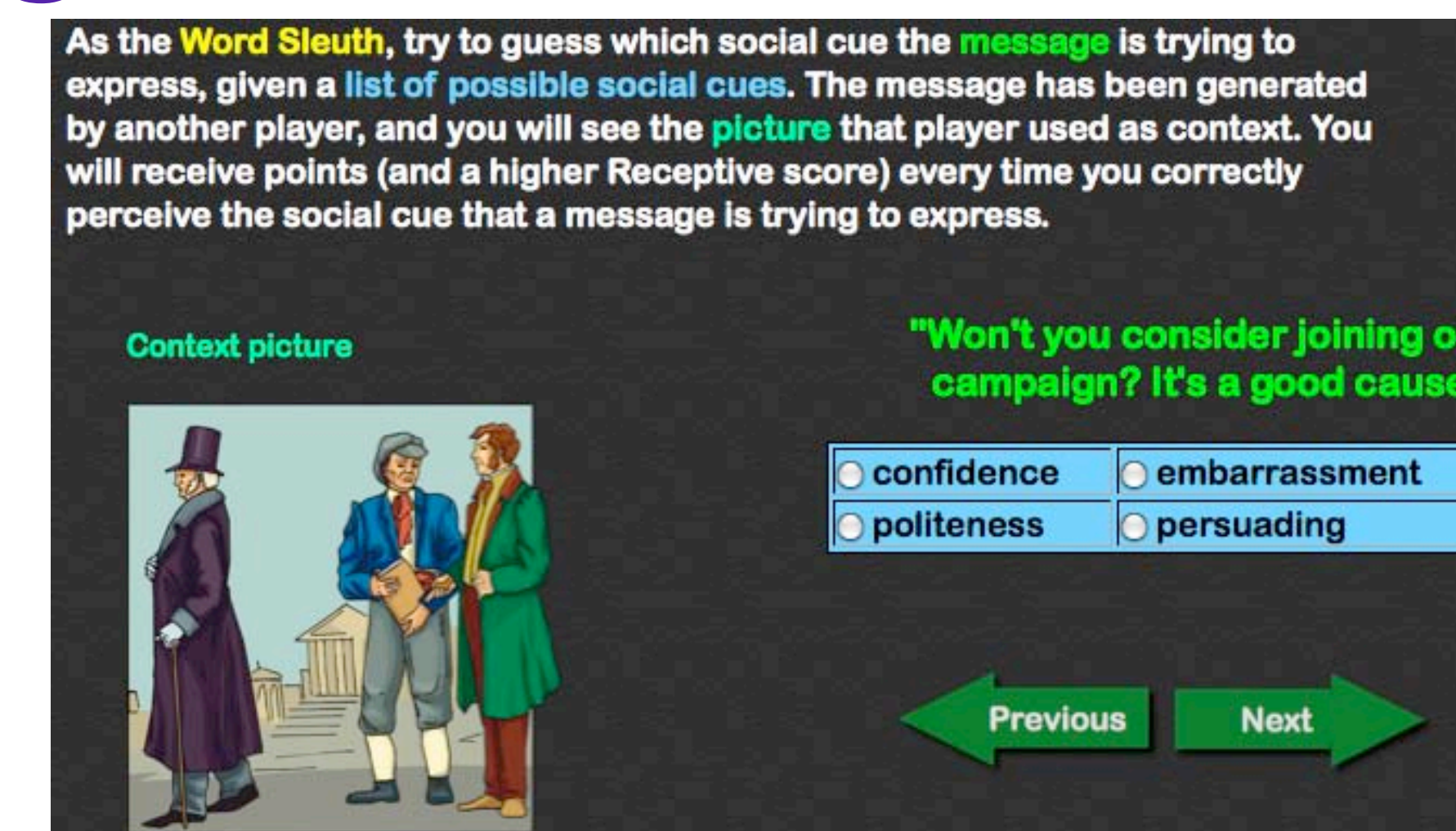
• Future Work: The GWAP methodology is easy to extend to create large-scale databases in both English and other languages, via online versions of the GWAP that run as web applications.

GWAP Design



Participants:
Game players

Motivation:
Getting points in the game for both generating messages that are easy to label and correctly labeling previously generated messages.



- **Social information types explored:** politeness, rudeness, embarrassment, formality, persuading, deception, confidence, disbelief
- **Asynchronous game play**
- **Context pictures:** Generic context pictures randomized, could be used for any prompt
- **Labeling:** messages labeled by multiple participants, only one label per message, participants asked to label more often than generate
- **Taboo words:** morphological variants of the social information type

GWAP Results

Participants:
58 English-speaking adults
(a mix of undergraduate students, graduate students, the authors, friends of the students, and friends of the authors, in order to simulate the varied mix of participants in an online GWAP; undergraduate students were compensated with course credit)

Messages Created: 1176
Message Labels Created: 3198

Sample Messages & Labels

Social Information Generated	Message
Labeled deception deception	“Oh yeah...your hair looks really great like that...yup, I love it...it, uh, really suits you...”
embarrassment embarrassment	“Oh... we're not dating. I would never date him... he's like a brother to me..”
disbelief disbelief	“Are you and him really friends?”
rudeness persuading	“James, Bree doesn't like you. She never did and never will!”
deception persuading	“I wasn't going to take anything from your storeroom, I swear! Really, I won't try to get inside again!”
politeness deception	“Your orange hair matches your sweater nicely”

How well do our non-experts do individually?

How well do they express the intended social information?
56% correctly expressed on average

How well do they perceive the intended social information?
58% correctly labeled on average

How reliable are the messages?

Confusion matrix for the human participants. The rows represent the intended social information for a message while the columns represent the labeled social information, averaged over messages and participants.

Diagonal represents correctly perceived messages.
Average correct: 57%.

Why so low? Some messages are better than others at expressing social information in a way obvious to humans.

What about the messages that the majority of humans agree on? **Is the crowd's perception better than individual perception?**

Confusion matrix for the human participants, where the majority of participants agreed on a message's intended social information and at least two participants labeled the message. The rows represent the intended social information for a message while the columns represent the labeled social information, averaged over messages and participants.

Diagonal represents correctly perceived messages.
Average correct: 71%

Wisdom of crowds effect observed

Remaining confusions:
• **deception** mistaken for **persuading** (.21) and **rudeness** (.10)
• **formality** mistaken for **politeness** (.34)

Individual-Labeled Messages

	deception	politeness	rudeness	embarrassment	confidence	disbelief	formality	persuading
deception	.37	.07	.10	.03	.09	.10	.04	.20
politeness	.05	.53	.05	.02	.03	.01	.20	.10
rudeness	.04	.01	.78	.02	.04	.04	.03	.03
embarrassment	.07	.09	.05	.56	.02	.13	.05	.03
confidence	.04	.04	.03	.01	.67	.05	.02	.13
disbelief	.10	.05	.05	.04	.07	.62	.02	.06
formality	.02	.34	.04	.02	.06	.03	.39	.10
persuading	.09	.06	.03	.01	.12	.03	.04	.61

Crowd-Labeled Messages

	deception	politeness	rudeness	embarrassment	confidence	disbelief	formality	persuading
deception	.45	.05	.10	.01	.07	.07	.03	.21
politeness	.03	.71	.03	.00	.01	.00	.13	.09
rudeness	.03	.00	.92	.00	.01	.02	.02	.00
embarrassment	.04	.08	.05	.69	.00	.11	.01	.02
confidence	.01	.04	.02	.01	.82	.01	.01	.09
disbelief	.05	.03	.02	.02	.05	.82	.00	.02
formality	.02	.34	.02	.01	.03	.03	.46	.10
persuading	.03	.05	.01	.00	.05	.03	.01	.82