Techniques for Emotion Classification

Julia Hirschberg

COMS 4995/6998

Thanks to Kaushal Lahankar

Papers for Today

- Cowie: “Describing the emotional states expressed in speech”, 2000
- Satoh et al: “Emotion clustering using the results of subjective opinion tests for emotion recognition in Infants’ cries”, 2007
- Nisimura et al: “Analyzing dialogue data for real-world emotional speech classification”, 2006

Cowie ’00: Describing the emotional states expressed in speech

- Emotions are not easily captured in words
- A suitable descriptive system for emotions does not exist as of yet.
- Objective - To encourage the speech community towards standardization of key terms and descriptive techniques.

What to Study?

- **Fullblown emotions**
  - Natural discrete units which can be counted and possess distinct boundaries
- **Emotional states**
  - An attribute of certain states. E.g. – “Her voice was tinged with emotion”.
- **Emotional systems**
- *How does each distinction influence research?*

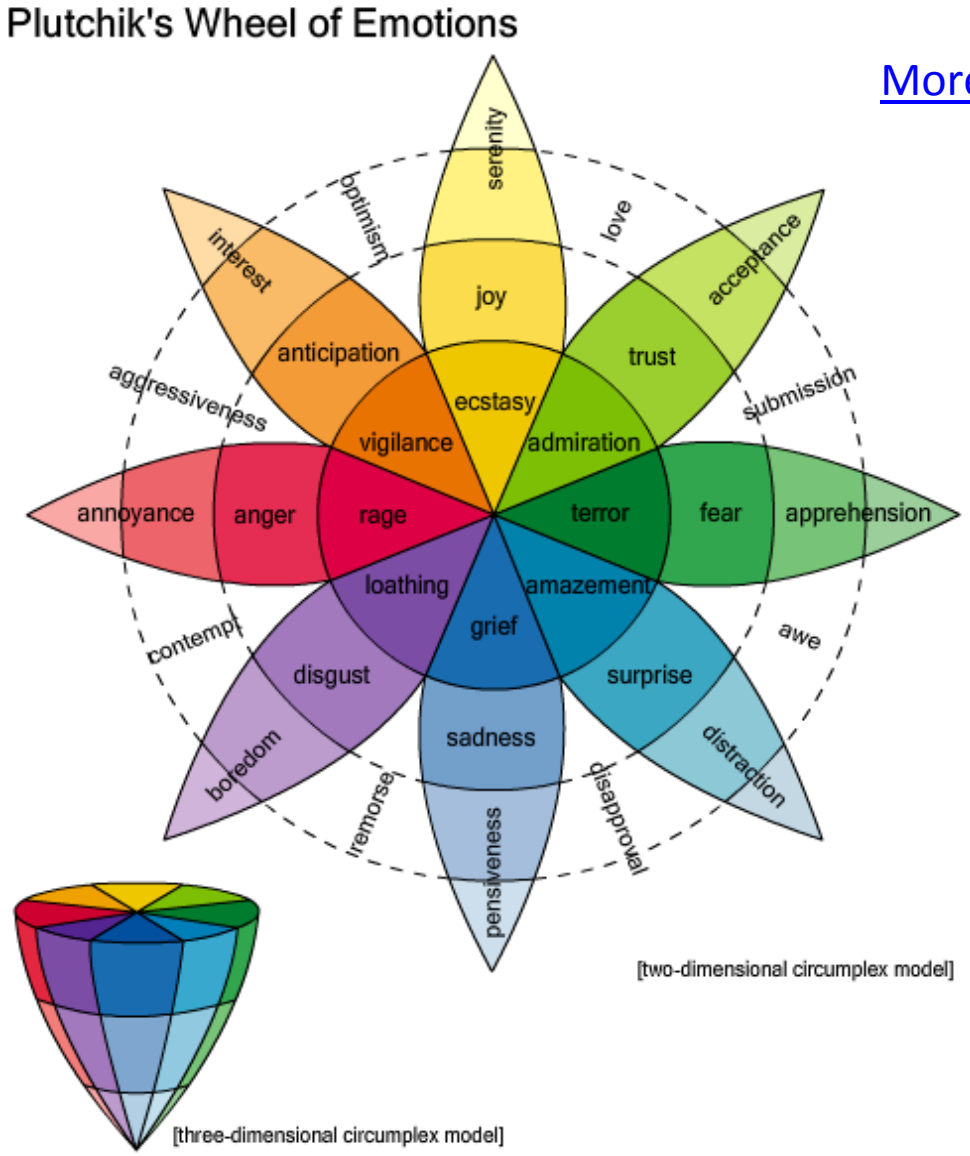
Cause and effect type descriptors

- Cause type
  - What are the emotion-related internal states and external factors causing speech to have particular characteristics?
  - Focus attention on physiological systems which can be used to describe emotions
- Effect type
  - Describes effect characteristics above would be likely to have on typical listener
  - Favor describing emotional states in terms of categories and dimensions that people find natural

Basic Emotion Categories

- **Basic (primary) emotions**
  - `Pure' emotional states
  - E.g. Ekman's Big 6: Fear, anger, happiness, sadness, surprise, disgust
- **Second order (derived) emotions**
  - Emotional states that are not so basic
  - E.g. chagrin, irritation

Wheel of emotions



More information

Other Representations

- Need to come up with a complementary representation that offers ways of drawing fewer, grosser distinctions to make things such as acquiring speech co-relates manageable
- *Feeltrace* might be such a system?

Emotion Related States

- ***Emotion-proper property*** :
  - Property ‘belonging to’ a type of emotion
  - Being frightening: emotion-proper property for fear
- Emotion terms surrounded by terms that resemble them but are different
  - Called *emotion related terms*
  - States associated with them called *emotion related states*

The states discussed

- Attitude
  - Categorization along an evaluative dimension
  - Appraisal
- Arousal
  - Degree of involvement

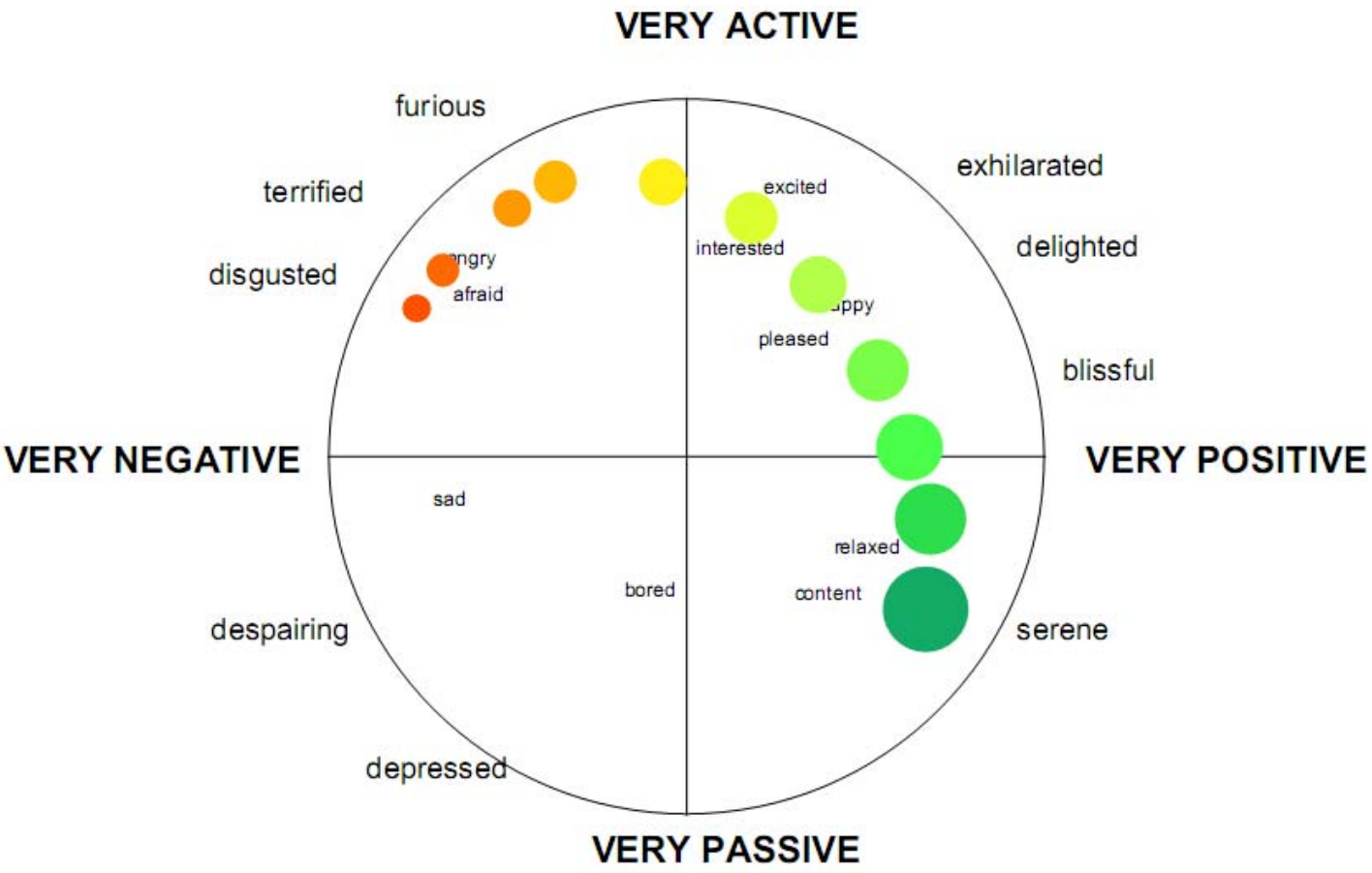
Biological Representations

- Assumption:
  - Descriptions of emotions are surrogates for descriptions of physiological states.
  - Can we replace verbal descriptions with physiological parameters?
- Which theory does this reflect?
- Pros and cons?

Continuous Representations

- Represent emotions in 2-D space in terms of Evaluation (X axis) and Activation (Y axis).
- Dimensions can be increased to represent the relationships which lie very close in the 2-D space.
- Pros and cons?

Feeltrace system



Structural Models

- Cognitive approach to describing emotions
- Based on hypothesis that distinct types of emotion correspond to distinct ways of appraising the situation evoking the emotion

Timing

- Short-lived or long-lived state
  - Sadness vs. grieving

How are Underlying Emotional Tendencies Expressed?

- Restraint
  - Display rules
- Ambivalence
- Humour
- Simulation

Tools for Describing Emotions

- Naturalistic Database – **Belfast Naturalistic Database** annotated using Feeltrace
- Feeltrace
- Basic Emotion Vocabulary
- Consider also
  - Whissell's Dictionary of Affect in Language (DAL): demo, revised version
  - Pennebaker's Linguistic Inquiry and Word Count (LIWC): demo, twitter analysis

Questions

Emotion clustering Using the results of subjective opinion tests for emotion recognition in infant's cries

- Objective – To design an emotion clustering algorithm for emotion detection in infants' cries
- Previous work
  - Acoustic analysis of an infant's cries has been performed.
  - Classification between "hunger" and "sleepiness" has been studied.
  - Some emotion detection products currently available in the market employ simple acoustic techniques.

Methodology

- Mothers asked to fill in an emotion table after recording each cry and to rate the intensity of each

Emotion	Anger	Sadness	Hunger	Surprise	……
Rank	0	4	2	0	……

Figure 1: *Example of the emotion table*

- Included pampered, anger, sdaness, fear, surprise, hunger, sleepiness, excretion, discomfort, painfulness
- Mothers could also see facial expression
- Baby-rearing experts did the same but only from recordings – agreement with mothers not high

Emotion Clustering

- Here, an emotion i was selected from a cluster X={e1,…,eI} and j, j≠i was selected from cluster Y={e1,…,e(i-1),eΦ,e(i+1),…eI}.

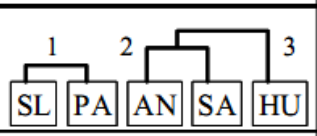
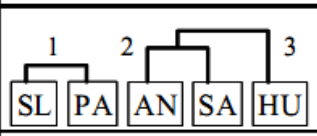
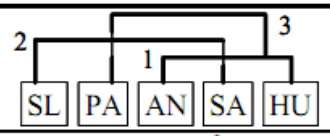
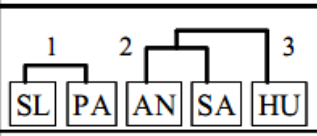
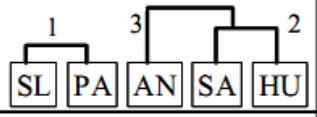
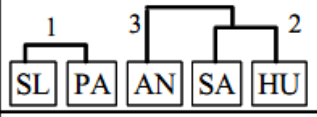
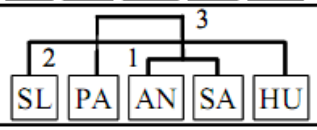
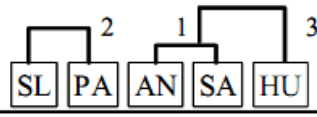
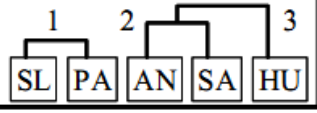
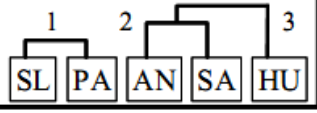
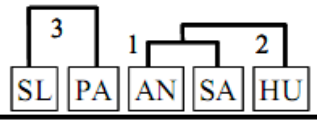
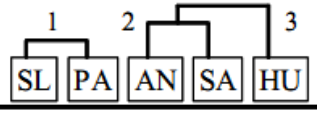
$$\hat{m},\hat{n} = \arg\min_{m,n\,(m\neq n)} H(Y \mid X)$$

$$= \arg\max_{m,n\,(m\neq n)} \Big\{ \sum_{i(i\neq m,n)} P(r)P(i \mid r)\log P(i \mid r) + P(r)P(\phi \mid r)\log P(\phi \mid r)$$

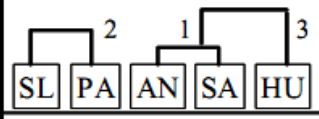
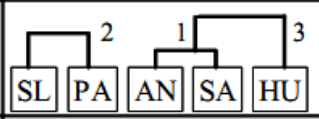
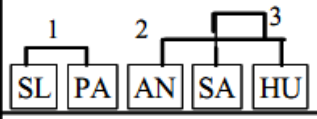
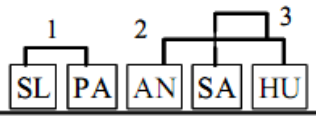
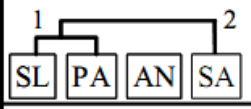
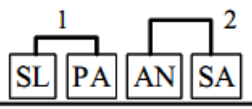
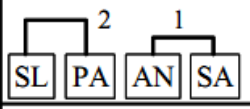
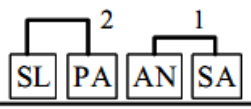
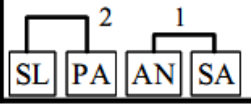
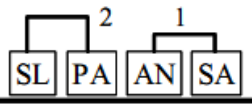
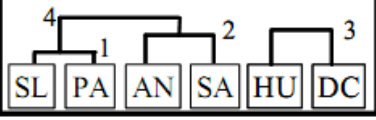
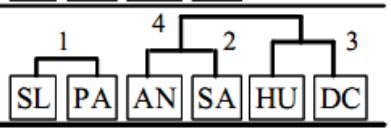
$$+ \sum_{i(i\neq m,n)} P(i)P(r \mid i)\log P(r \mid i) + \sum_{i(i\neq m,n)} \sum_{j(j\neq m,n)} P(i)P(j \mid i)\log P(j \mid i) \Big\} \quad (2)$$

- Form of hierarchical clustering where the conditional entropy is the objective function to be minimized.

Clustering trees

Rank	Three and more (R3)		Two and more (R2)	
Method	Method I	Method II	Method I	Method II
Mothers and experts (Set-M&D)				
Mothers (Set-M)				
Experts (Set-D)				

(2-a) Clustering trees using the results performed by infants' mothers, baby-rearing experts, and both of them

Rank	Three and more (R3)		Two and more (R2)	
Method	Method I	Method II	Method I	Method II
Infant A				
Infant B				
Infant C				

(2-b) Clustering trees for three infant sets using the results performed by infants' mothers and baby-rearing experts

Figure 2. *Obtained clustering trees using the results of subjective opinion tests* (SL: sleepiness, PA: pampered, AN: anger, SA: sadness, HU: hunger, DC: discomfort)

Emotion Recognition

- Find the emotion given the segment sequence of a cry
- Determine the probability of the acoustic evidence q occurring when the infant utters the sequence z with the emotion e (argmax(P(q|e,z)_
  - Acoustic models for each segment for each emotion cluster
  - Train HMMs and use to recognize cries
  - Results with 2 clusters: 75% accuracy

Nisimura et al '06: Analyzing Dialogue Data for Real-World Emotional Speech Classification

- Goal: Classify child utterances wrt 'delightful' or 'hateable' emotions
- How did they obtain and annotate the data?
- What was the method of analysis?
  - Factor analysis
  - Support vector machine
  - Regression analysis
  - Jackknife method
- What can we conclude from this experiment?

Discussion

- Are 10 emotions actually needed?
- Instead of relying on the input from baby-rearing experts, could the inputs from the mothers have been used?
- Is the input data reliable?
- Improvements?