The Ordinal Nature of Emotions

Georgios N. Yannakakis, Roddy Cowie and Carlos Busso

The story

"It seems that a **rank**-based *FeelTrace* yields higher inter-rater agreement…"

"Indeed, *FeelTrace* should actually be used this way…
(!) Go talk to Carlos; see you in two years… bye!..."

This paper

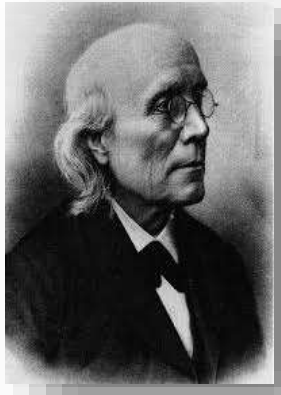
A **thesis**: emotions are intrinsically **ordinal** (relative)
…and the benefits of representing them that way are many!

Our thesis is supported by **theoretical arguments** across disciplines and **empirical evidence** in Affective Computing

Our wish: **reframe** the way emotions are viewed, represented and analysed computationally

The Background (Psychology)

One of the first challenges in Psychology



Mapping the intensities of responses to particular stimuli

That is basic to affective computing: we call it **labelling**

Two approaches have a long history

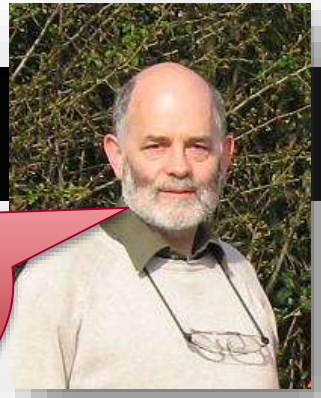
- The older (Fechner) was based on comparing stimuli, and finding '**just noticeable differences**'
- Much later, Stevens introduced '**magnitude estimation**' – asking people to give a number. Twenty years ago, psychologists tried a magnitude estimation approach to labelling. The data are in, and we know it doesn't work as straightforwardly as they hoped.

The core finding is simple…

When raters are presented with a *piece of data* and asked to assign a *magnitude* describing an emotional response, they tend to **disagree** quite substantially.

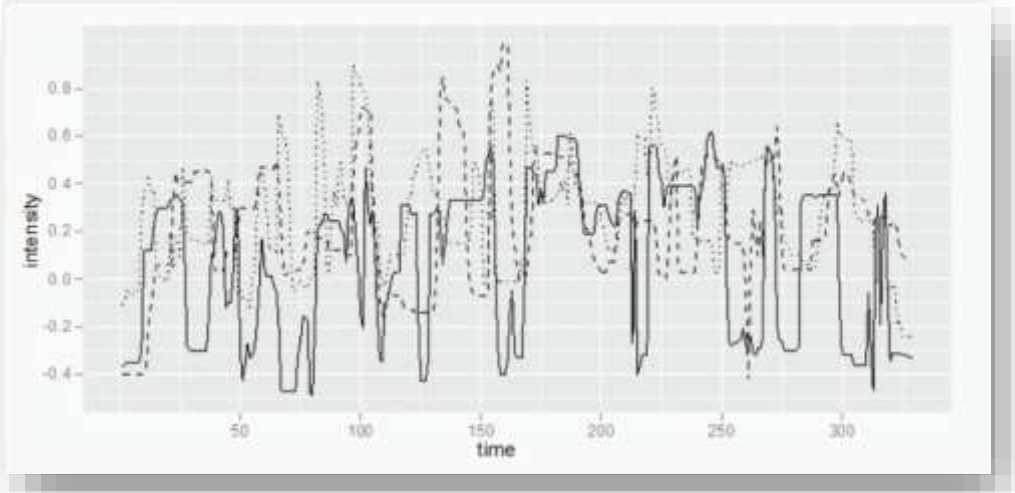
Douglas-Cowie et al. "**Multimodal databases of everyday emotion: Facing up to complexity**," *Ninth European Conference on Speech Communication and Technology*. 2005.

The core finding is simple…



No point in modelling noise

- That is not a criticism of the **constructs** used, like valence and arousal
- Sometimes agreement is quite good
  - But not often enough



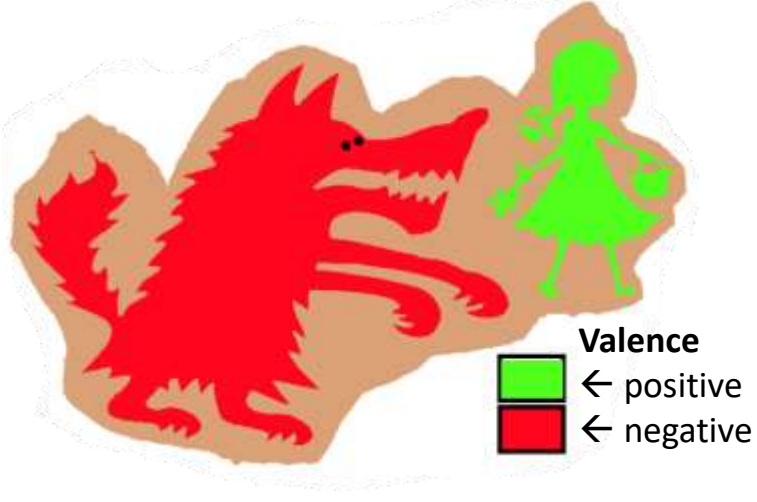
Cowie et al., "**Tracing emotion: an overview**" *International Journal of Synthetic Emotions,* 2012

The reasons are obvious and long known…

…we just did not know how serious they would be…

Reason 1 Data are typically multivalued

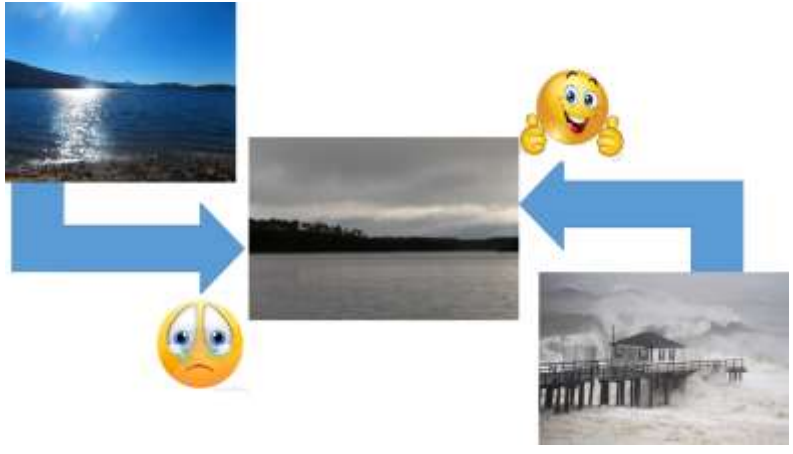
- A scene will contain *multiple elements*, which have *different valences*, and there is no self-evident way to reduce them to a *single measure*.



The reasons are obvious and long known…

Reason 2 Adaptation level

- Say today is a **grey day** (obviously in Belfast); what feelings will it evoke?
  - **–ve:** if it's ending a *sunny spell*
  - **+ve:** if we are coming out of a *hurricane*



But labelling is associating a value; So, which should we associate?

The Background (Beyond Psychology)

Marketing

- It seems that **societal** or **ethical** values are acquired, internalized and organized in a *hierarchical manner*. The **ranking** approach naturally helps respondents to discover, reveal and crystallize that hierarchy
- The empirical evidence is **strong**: ranks are more effective (than ratings) at reducing response biases in cross-cultural settings

Johnson et al., "**The relation between culture and response styles: Evidence from 19 countries**," *Journal of Cross-cultural psychology*, vol. 36, no. 2, pp. 264–277, 2005

Neuroscience

- Each time we are presented with a stimulus, we construct and store an **anchor** (or **somatic marker**)
- We use somatic markers as drivers for making choices
- Affect guides our attention towards preferred options and, in turn, simplifies the decision process for us!

Further evidence (in monkeys and humans) suggests that our brain **encodes** values in a **relative** fashion

Damasio, "**Descartes' error: Emotion, rationality and the human brain**," 1994.

Seymour and McClure, "**Anchors, scales and the relative coding of value in the brain**," *Current opinion in neurobiology*, 2008

Behavioural Economics

"...it is safe to assume that **changes** are more accessible than **absolute** values…"

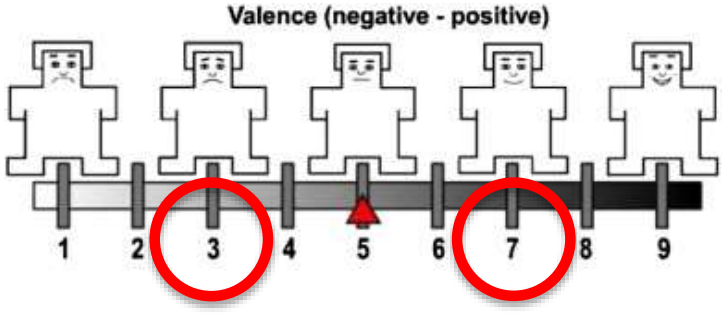
Daniel Kahneman. **A perspective on judgment and choice: mapping bounded rationality**. *American psychologist*, 58(9):697, 2003

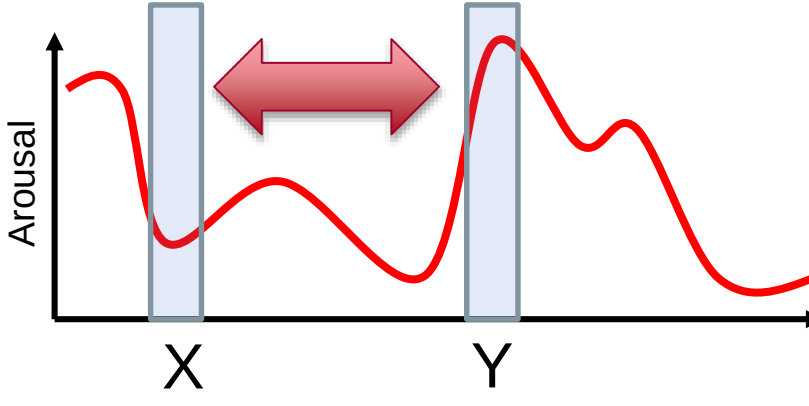
AI and Machine Learning

- **Preference learning** is inspired by and built upon humans' limited ability to express their preferences *directly* in terms of a specific (subjective) value function
- Our inability is mainly due to the
  - **subjective nature** of a preference
  - **cognitive load** for assigning specific values to each one of the options
- It is more **natural** to express preferences about a number of options; and this is what we end up doing normally.

S. Kaci, **Working with preferences: Less is more**. Springer Science & Business Media, 2011.

Summary: relationships matter… not their magnitude

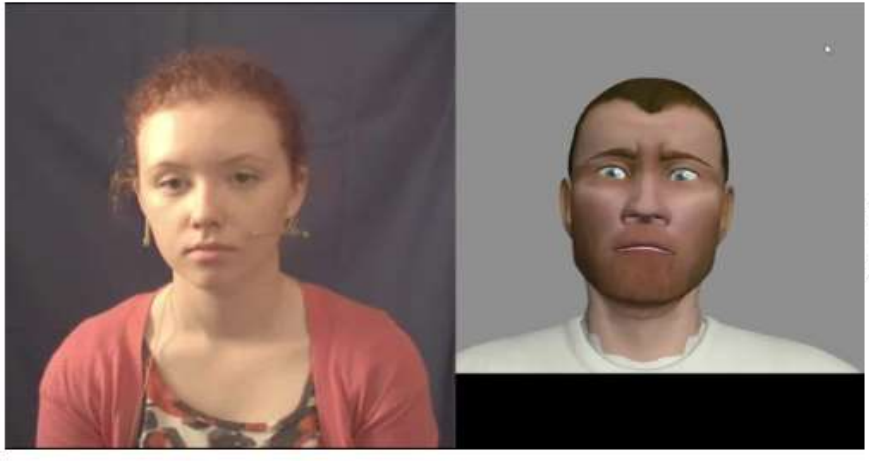


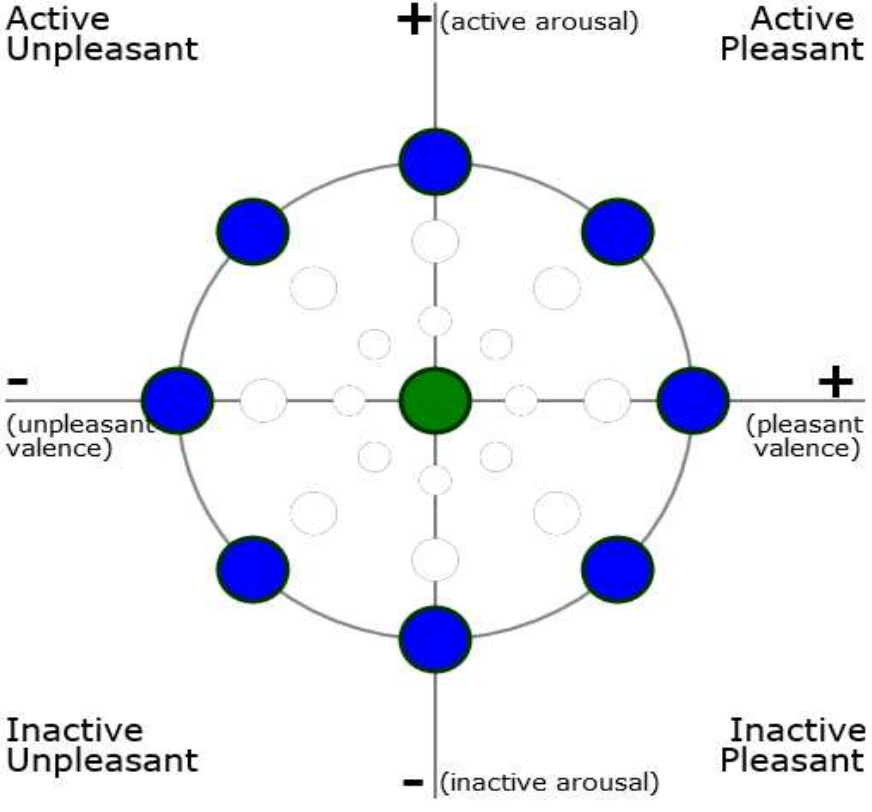


The Evidence

Video Annotation: *AffectRank*

Yannakakis and Martinez, **Grounding Truth via Ordinal Annotation,** *Affective Computing and Intelligent Interaction*, 2015.



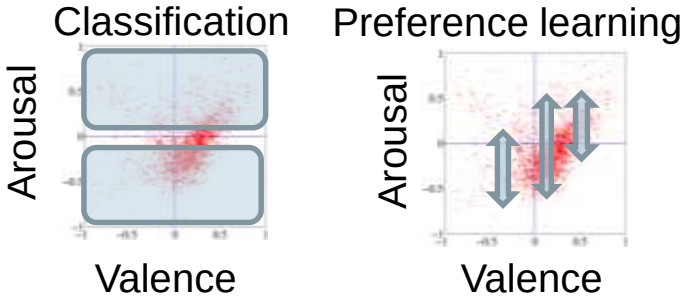


Available at:

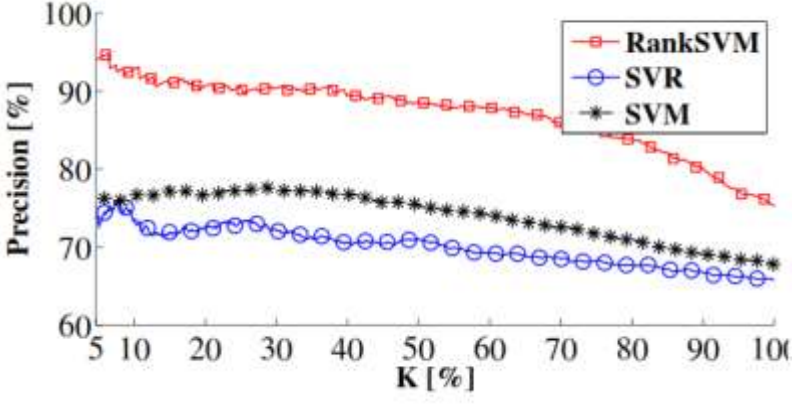
https://github.com/TAPeri/AffectRank

Speech: Preference Learning For Emotion Recognition

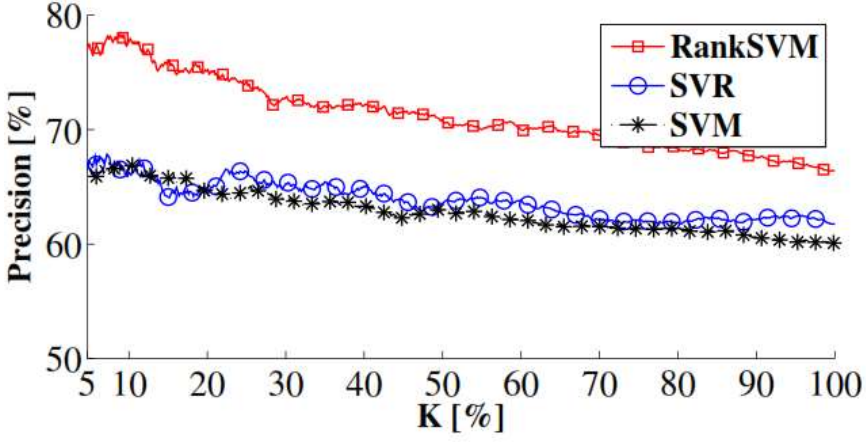
Lotfian and Busso, "**Practical considerations on the use of preference learning for ranking emotional speech**," in IEEE ICASSP 2016



- Better use of the corpus:
  - n(n-1)/2 potential pairs for training
- More reliable labels
- Better performance (precision@K)



Arousal

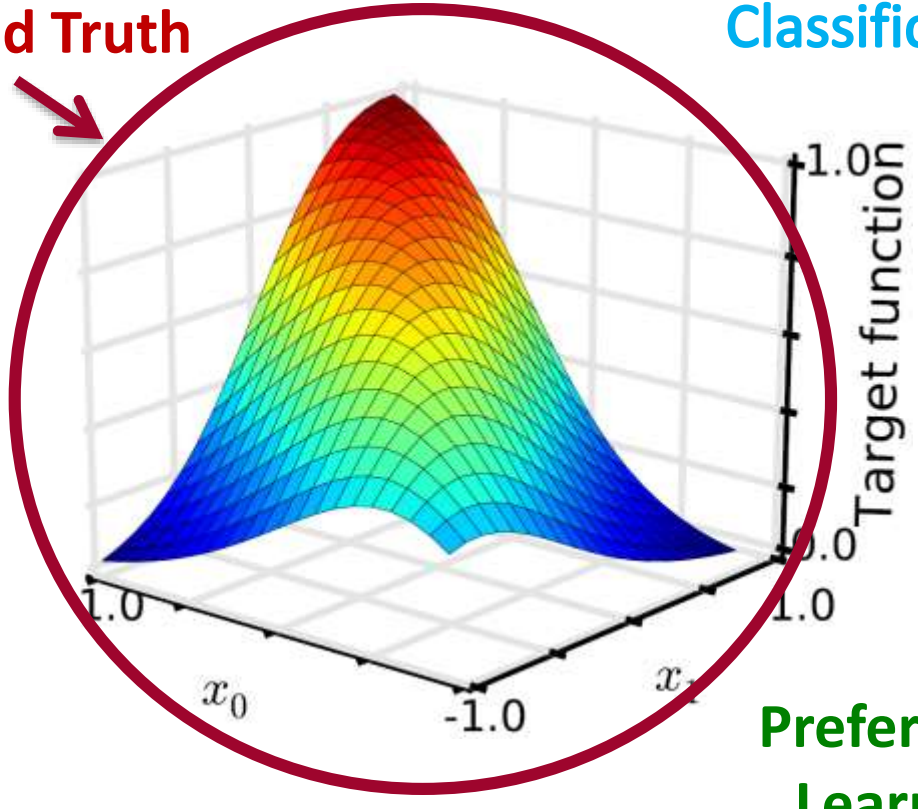


Valence

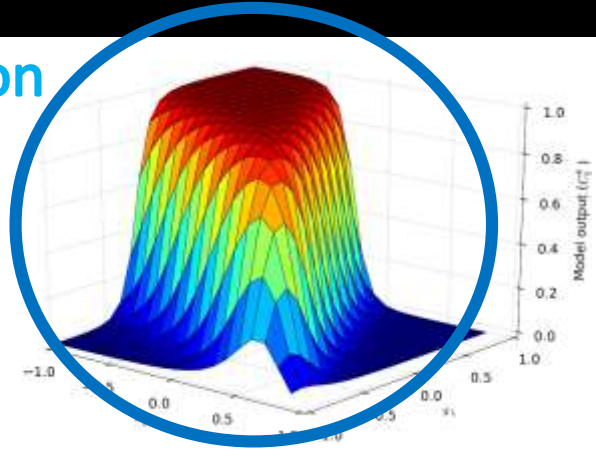
Speech and Games: Classes vs Preferences

Martinez, Yannakakis and Hallam, **Don't classify ratings of affect; Rank them!** *IEEE Trans. on Affective Computing*, 2014.

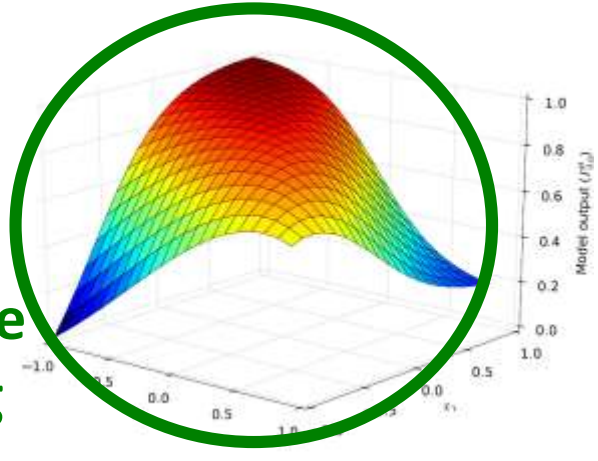
Ground Truth



Classification



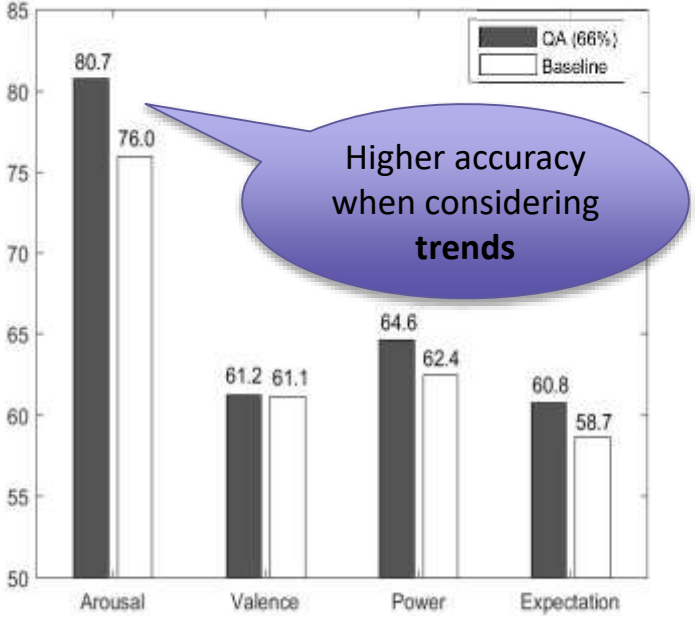
Preference Learning

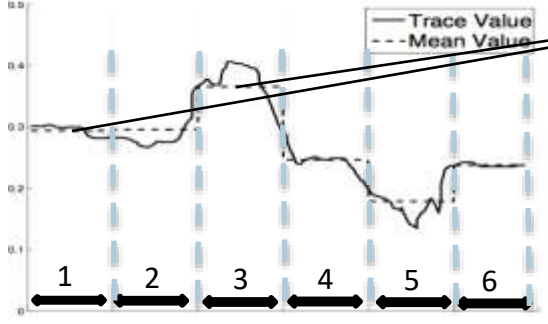


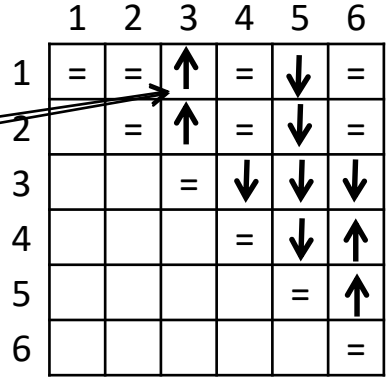
Speech Annotation: Qualitative Agreement Analysis

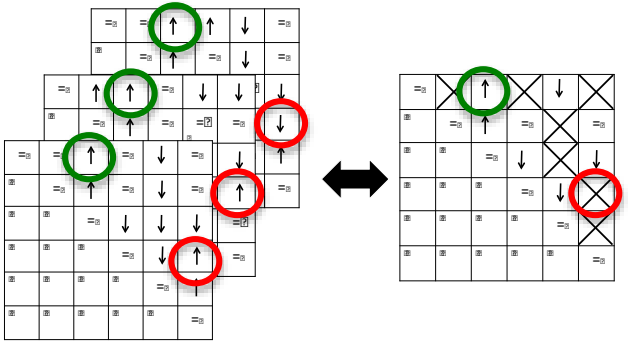
Parthasarathy et al., "**Using agreement on direction of change to build rank-based emotion classifiers**," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2016.

- Divide trace into bins
- Look for trends
- Create **preference learning** models based on the trends







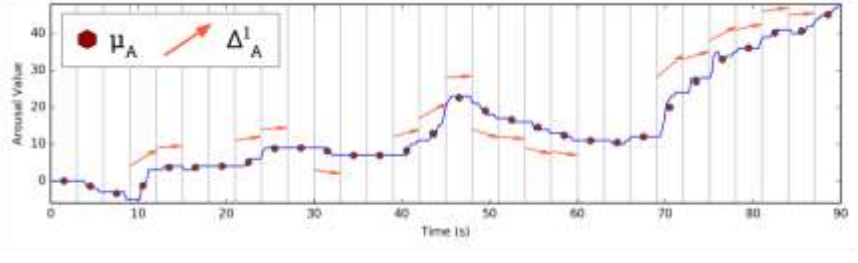


Video Annotation: *RankTrace*

Lopes, Liapis, and Yannakakis, **RankTrace: Relative and Unbounded Affect Annotation** *ACII*, 2017
Camilleri, Yannakakis and Liapis, **Towards General Models of Player Affect**, *ACII*, 2017

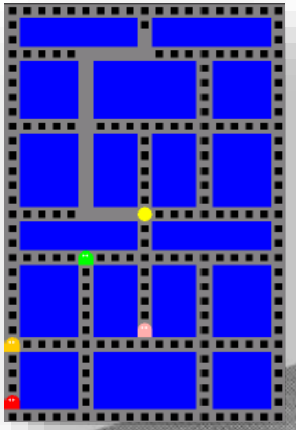
- Better predictors of ground truth
- More **general** affect models across tasks

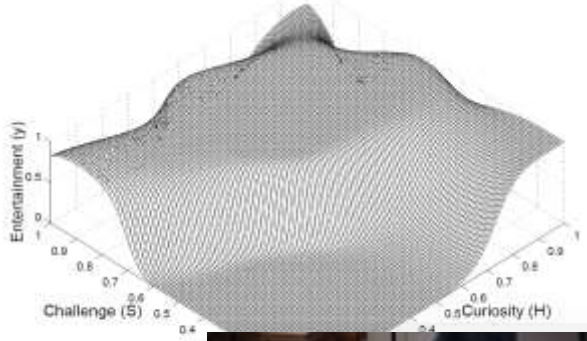
Available@emotion-research.net

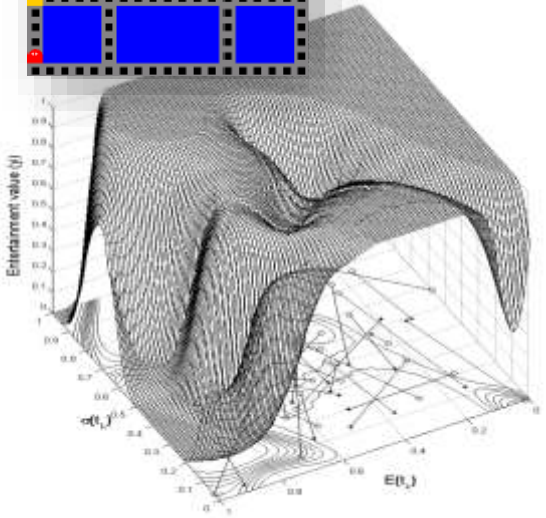


Games: Ratings (Likert) vs Preferences (Ranks)

Yannakakis and Hallam, **Rating vs. Preference: A comparative study of self-reporting**, ACII, 2011
Yannakakis and Martinez, **Ratings are Overrated!** *Frontiers in Human-Media Interaction,* 2015







X is engaging

Disagree Agree

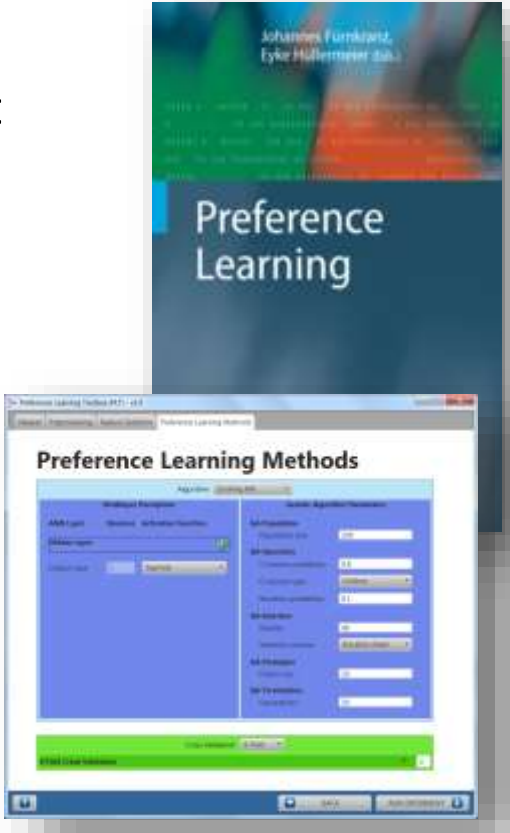
-2 -1 0 1 2

- [ ] X is **more/less** engaging than Y
- [ ] Both are **equally** engaging
- [ ] **Neither** is engaging

So I have Ranks; What’s Next?

Preference Learning for Affective Computing

- Tutorial: ACII 2009, Amsterdam
- An approach with growing interest since then for affect detection and retrieval through images, videos, music, sounds, speech, games, and text
- Several PL algorithms available.
  - SVM (RankSVM)
  - Shallow and Deep Neural Networks
  - Gaussian Processes
  - …
  - Some of them in the PL Toolbox (emotion-research.net)
- Domains: healthcare, education, entertainment, art,…

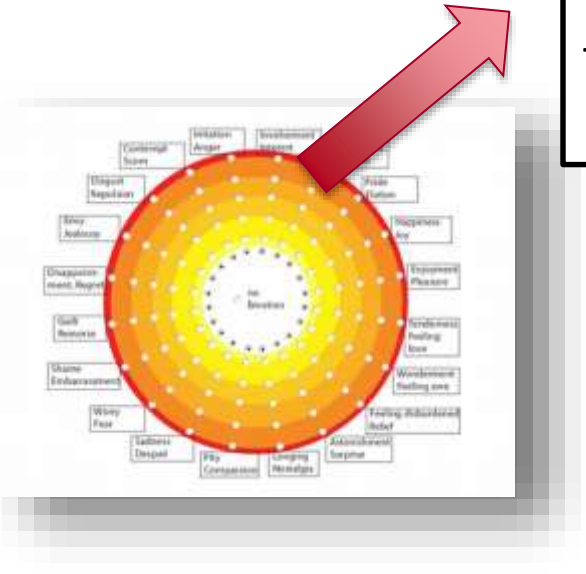


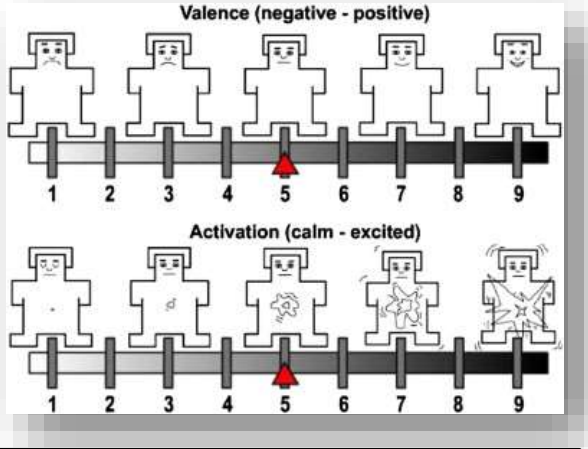
G. N. Yannakakis, "**Preference learning for affective modeling**," in *Affective Computing and Intelligent Interaction,* 2009

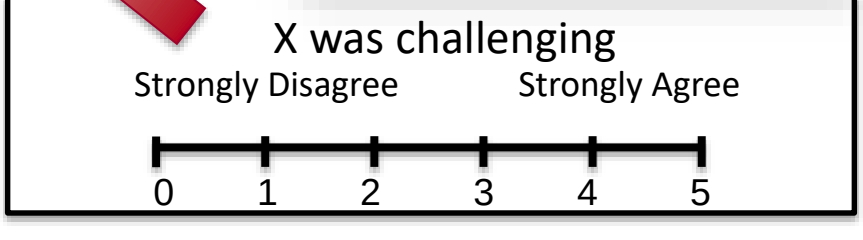
What if Ranks are not Available?

Martinez, Yannakakis and Hallam, **Don't classify ratings of affect; Rank them!** *IEEE Trans. on Affective Computing*, 2014.

X is **more/less** challenging / frustrating / arousing / boring / fearful / … than **Y**







The Criticism

The Criticism and our Response

"More information (i.e. intensity) is always good to have.."

- Less is more! Intensity is actually maintained (it is lying under the preference). More information biases the model

"More options are required in ranks; one stimulus is not enough…"

- This is their very strength! Our anchor/marker/reference is not retrieved unconsciously or intuitively; it is forced! Our **reference** is a **real option** we use during the annotation.

"Analysis is harder with ordinal data…"

- Multiple data visualization and processing techniques are available nowadays: classical correlation analysis to statistical significance tests to modern ML approaches

Takeaway

- Our thesis is not new… but it **reframes** AC
- We are not alone… but we hope more will join the **ordinal** stance
- The evidence keeps coming…
- It seems that we best encode subjective values in **relative** terms
- Machine learning should probably do so too!
- **Preference learning** is a way forward!
- Benefits: reliability, validity, generality

Thank you!