Capturing Ambiguity in Crowdsourcing Frame Disambiguation

Anca Dumitrache, Lora Aroyo, Chris Welty

VU UNIVERSITY AMSTERDAM

What is FrameNet?

FrameNet: computational linguistics resource based on the frame semantics theory (Baker, Fillmore, Lowe, 1998)

- collection of **semantic frames**
- **documents** annotated with these frames

semantic frame: abstract representation of a word sense, describing a type of *entity*, *relation*, or *event* grounded in *roles* implied by the frame

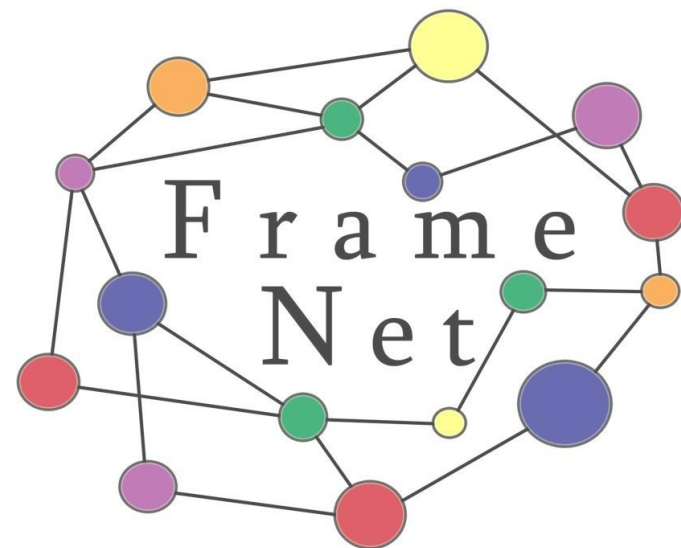e.g. *from* & *to* are roles in a *movement* frame

Anca Dumitrache
@anca_dmtrch
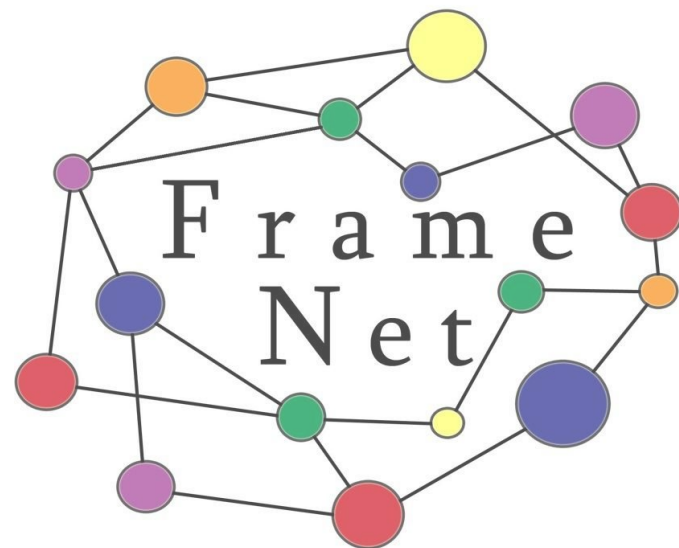
Frame Disambiguation

= task of selecting the best frame for a word phrase

> We try to avoid shops ***aimed*** at tourists.
>
> - purpose (*)
> - aiming
>
> *The frame picked by the expert is marked with (*).*

Frame Disambiguation

= task of selecting the best frame for a word phrase

Illegal ***skimming*** of profits is rampant.

- removing
- theft
- commiting crime
- cause change
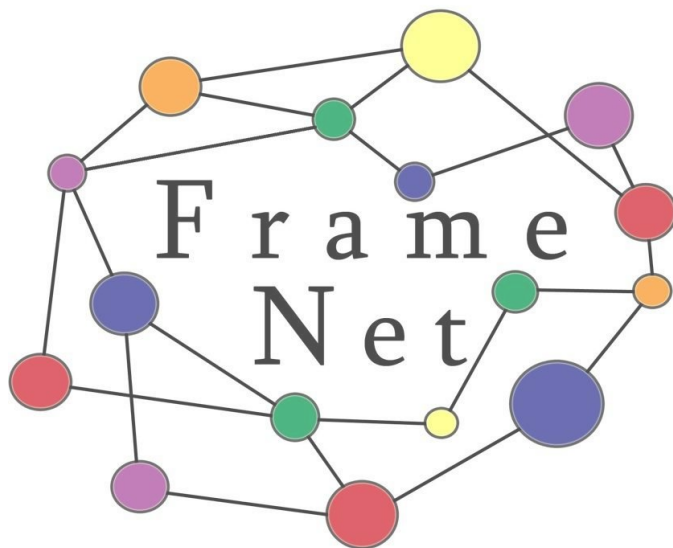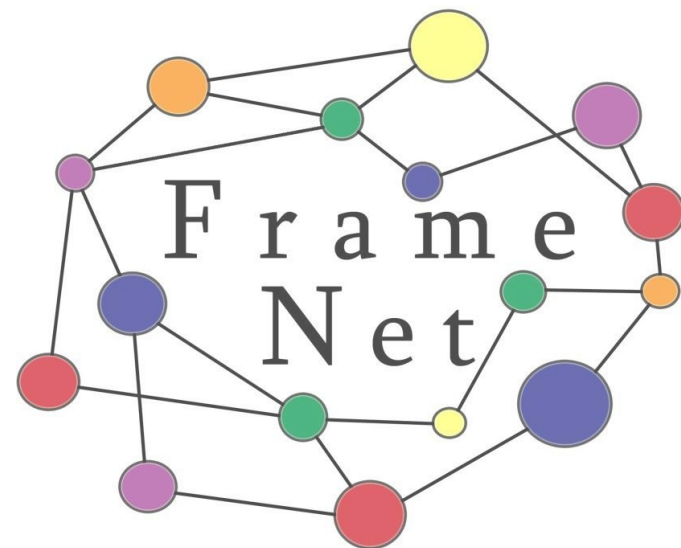
Anca Dumitrache
@anca_dmtrch

Frame Disambiguation

= task of selecting the best frame for a word phrase

Illegal ***skimming*** of profits is rampant.

- removing (*)
- theft
- commiting crime
- cause change

The frame picked by the expert is marked with (*).

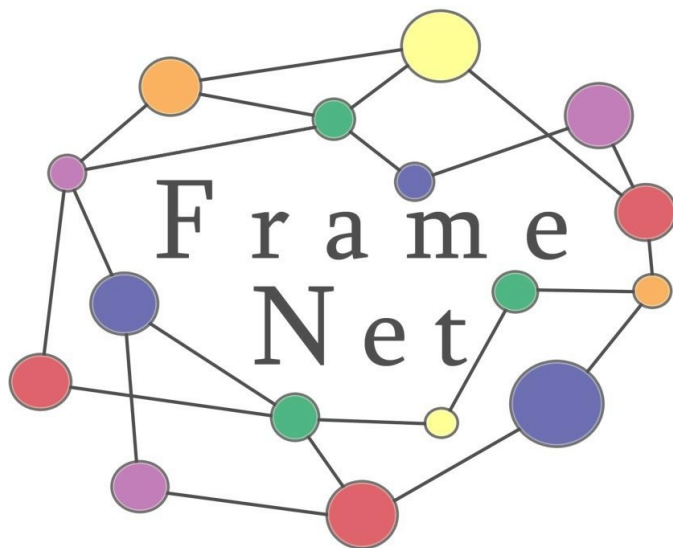**What does the crowd think?**

Anca Dumitrache
@anca_dmtrch

Frame Disambiguation

= task of selecting the best frame for a word phrase

Illegal *skimming* of profits is rampant.

- removing (*) **→ 7 votes**
- theft **→ 6 votes**
- commiting crime **→ 6 votes**
- cause change **→ 4 votes**

The frame picked by the expert is marked with (*).

Anca Dumitrache
@anca_dmtrch

Dataset

- **433 sentence-word pairs** from FrameNet corpus
  - words have *expert annotation*
  - POS: verb, noun

- Pre-processing to find **candidate frames for each word**:

  - match word to *synonym sets* in WordNet corpus (Miller, 1995)

  - match synonym set to FrameNet frame using *Framester* corpus (Gangemi et al., 2016)

Crowdsourcing task

The sentence:

Anarchism is a political philosophy that **advocates** self-governed societies based on voluntary institutions.

What are the possible meaning(s) of advocates in the context of the sentence above? Check ALL that apply.

☐ **Communication:** A *Communicator* conveys a *Message* to an *Addressee*; the *Topic* and *Medium* of the communication also may be expressed.

Frame definition

> Click to hide examples where the highlighted word expresses **Communication**

It **says** a lot that he didn't come back.
Putting his arm around her protectively achieved nothing but **announcing** to their captors their vulnerability.
This painting really **speaks** to me.

Example sentences for each frame, toggled by button

Multiple choice task

☐ **Attempt suasion:** The *Speaker* expresses through language his wish to get the *Addressee* to act. There is no implication that the *Addressee* forms an intention to act, let alone acts.
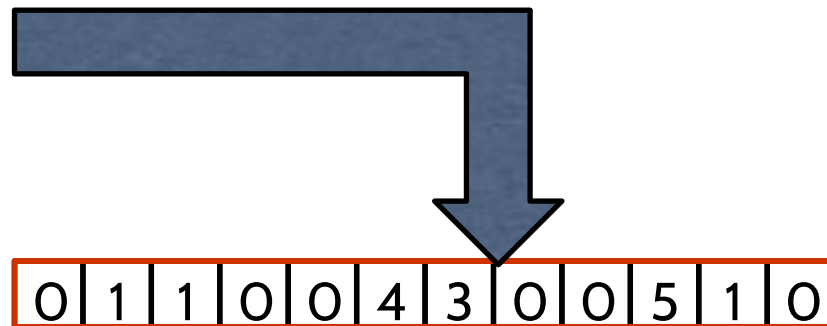
Frame definition

> Click to see examples where the highlighted word expresses **Attempt suasion**

Anca Dumitrache
@anca_dmtrch

Worker Vectors

	Communication	Attempt suasion	Cause change	. . .						
W1:					1	1			1	
W2:		1							1	
W3:					1					
W4:				1					1	
W5:	1								1	
W6:				1	1					
W7:								1		
W8:										1

0	1	1	0	0	4	3	0	0	5	1	0

Sentence Vector

Anca Dumitrache
@anca_dmtrch

CrowdTruth metrics

Frame-Sentence Score (FSS): the degree with which a particular frame matches the sense of the word in the sentence

$$FSS(s, f) = \frac{\sum \text{workers that picked } f \text{ in } s \cdot \text{worker quality}}{\sum \text{workers for } s \cdot \text{worker quality}}$$

Sentence Quality Score (SQS): overall worker agreement over one sentence, measured with cosine similarity

$$SQS(s) = \frac{\sum_{w1,w2 \in \text{ workers for } s} \text{weighted cos sim}(\vec{w1}, \vec{w2}, FQS) \cdot \text{ worker quality}(w1, w1)}{\sum_{w1,w2 \in \text{workers for } s} \text{worker quality}(w1, w2)}$$

Frame Quality Score (FQS): agreement over a frame in all sentences where the frame was picked at least once

$$FQS(f) = \frac{\sum_{FSS(s,f)>0} FSS(s, f)SQS(s)}{\sum_{FSS(s,f)>0} SQS(s)}$$

Crowd vs. FrameNet experts ground truth
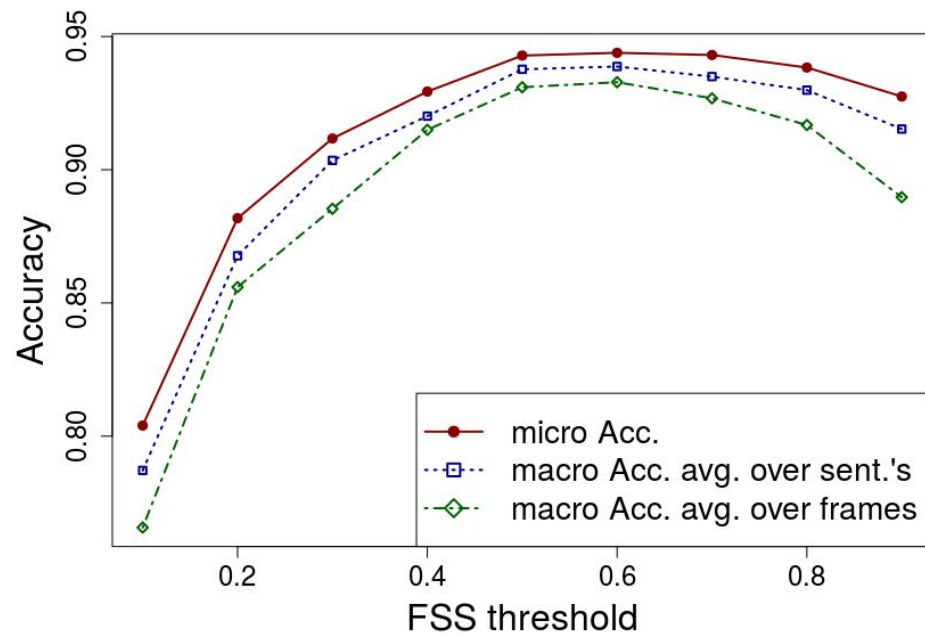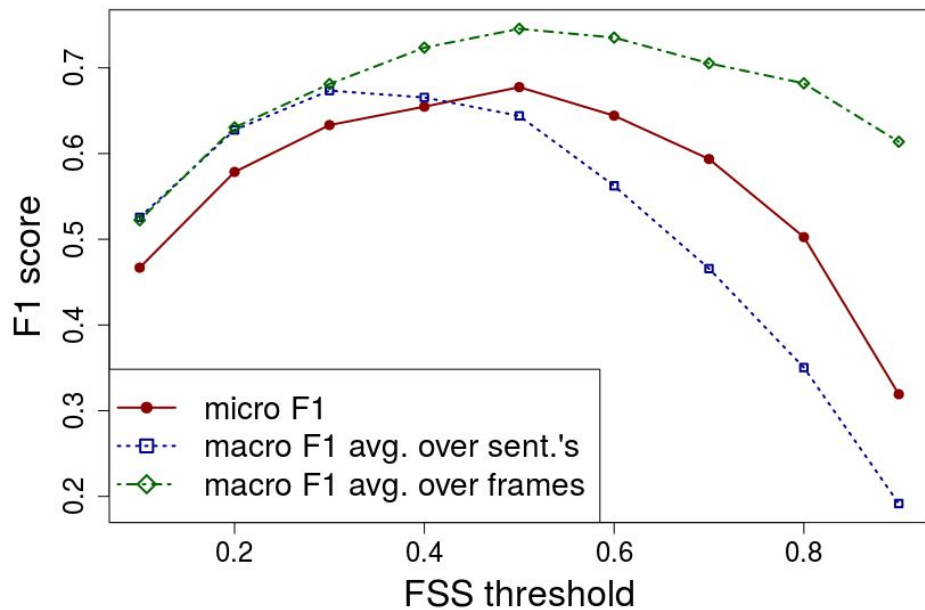
Crowd performance is comparable to the experts.

Anca Dumitrache
@anca_dmtrch

When crowd & expert disagree

Crowd **misunderstood** the frame definition.

Anca Dumitrache
@anca_dmtrch

When crowd & expert disagree

Crowd **misunderstood** the frame definition.

Information in the sentence is **incomplete**.

> The *investigation* has been stymied, stopped, obstructions thrown every step of the way.

Crowd: criminal investigation (FSS = 0.804)

Expert: scrutiny (FSS = 0.305)

When crowd & expert disagree

Crowd **misunderstood** the frame definition.

Information in the sentence is **incomplete**.

The *investigation* has been stymied, stopped, obstructions thrown every step of the way.

Crowd: criminal investigation (FSS = 0.804)

Expert: scrutiny (FSS = 0.305)

Crowd is **correct**.

Does *supersizing* cause obesity?

Crowd: cause to start (FSS = 0.804)

Expert: causation (FSS = 0.608)

Crowd still picked the expert frame, but with lower FSS.

FSS as a measure of how clearly the frame is expressed in the sentence

Example sentences with *removing* frame:

The frame picked by the expert is marked with *.

Egypt has provided no evidence demonstrating the *elimination* of its biological weapons.

removing* - **FSS = 0.938**
cause change - FSS = 0.175

FSS as a measure of how clearly the frame is expressed in the sentence

Example sentences with *removing* frame:

The frame picked by the expert is marked with *.

Egypt has provided no evidence demonstrating the *elimination* of its biological weapons.

> *removing** - FSS = 0.938*
> cause change - FSS = 0.175

The Syrian Mujahiddin asked Hussein to *overthrow* the regime of Hafiz Al Assad.

> change of leadership* - FSS = 0.847
> *removing - FSS = 0.539*

Anca Dumitrache
@anca_dmtrch

FSS as a measure of how clearly the frame is expressed in the sentence

Example sentences with *removing* frame:

The frame picked by the expert is marked with *.

Egypt has provided no evidence demonstrating the *elimination* of its biological weapons.

> *removing* * - *FSS = 0.938*
> cause change - FSS = 0.175

The Syrian Mujahiddin asked Hussein to *overthrow* the regime of Hafiz Al Assad.

> change of leadership * - FSS = 0.847
> *removing* - *FSS = 0.539*

Illegal *skimming* of profits is rampant.

> *removing* * - *FSS = 0.532*
> theft - FSS = 0.494
> commiting crime - FSS = 0.459
> misdeed - FSS = 0.431
> cause change - FSS = 0.273

Anca Dumitrache
@anca_dmtrch

SQS as a measure of how ambiguous the sentence is

Example sentences with *removing* frame:

The frame picked by the expert is marked with *.

Egypt has provided no evidence demonstrating the *elimination* of its biological weapons.

removing[*] - FSS = 0.938
cause change - FSS = 0.175

SQS = 0.841

The Syrian Mujahiddin asked Hussein to *overthrow* the regime of Hafiz Al Assad.

change of leadership[*] - FSS = 0.847
removing - FSS = 0.539

SQS = 0.669

Illegal *skimming* of profits is rampant.

removing[*] - FSS = 0.532
theft - FSS = 0.494
commiting crime - FSS = 0.459
misdeed - FSS = 0.431
cause change - FSS = 0.273

SQS = 0.366

Anca Dumitrache
@anca_dmtrch

FQS as a measure of how ambiguous the frame is

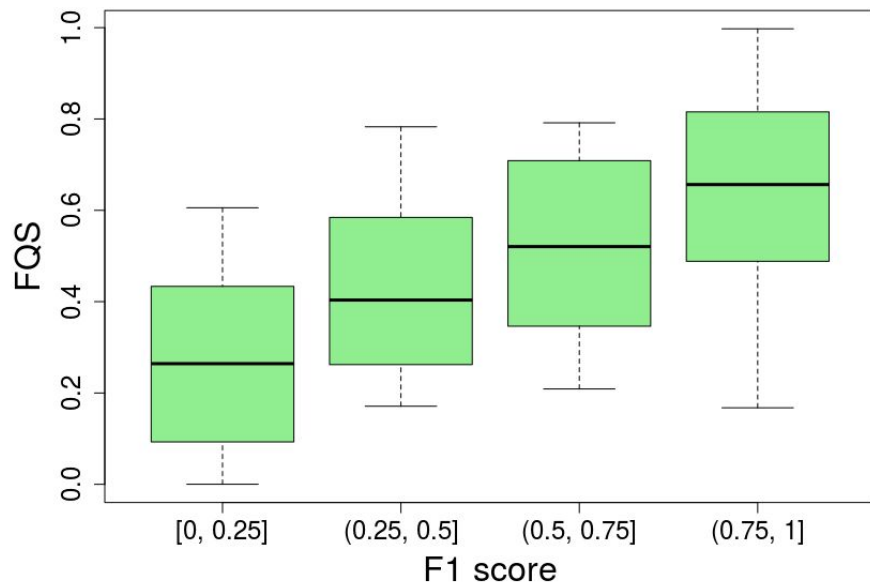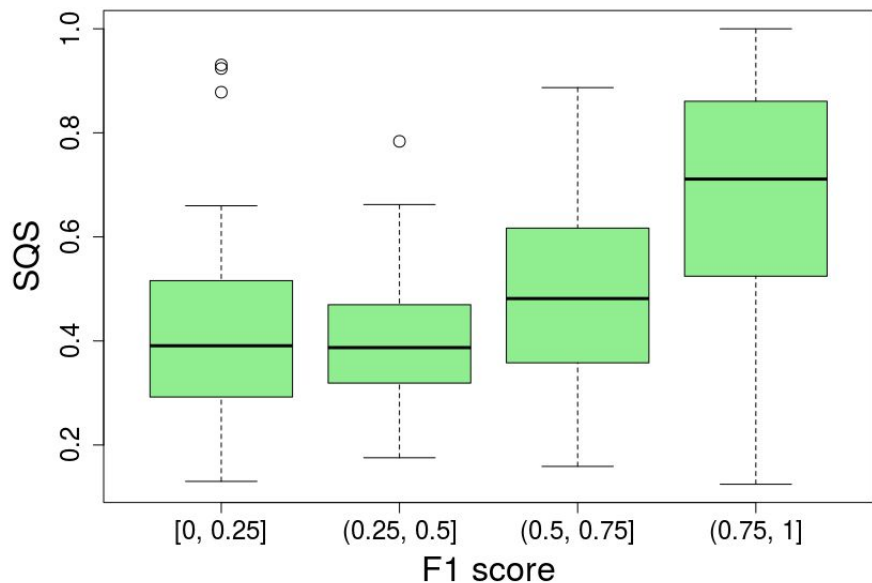**Concrete frames** have high FQS.

e.g. *removing*

Abstract frames have low FQS.

e.g. *cause change*

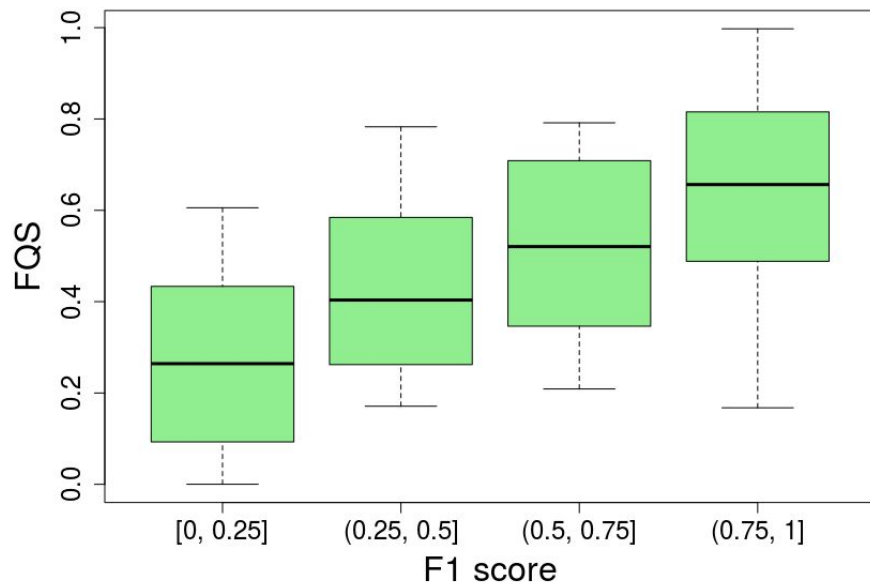Frames with **overlapping definitions** have low FQS.

e.g. *objective influence* & *subjective influence*

Anca Dumitrache
@anca_dmtrch

SQS and FQS vs. Expert ground truth

When the crowd workers agree with each other, they also agree with the expert.

Anca Dumitrache
@anca_dmtrch

SQS and FQS vs. Expert ground truth
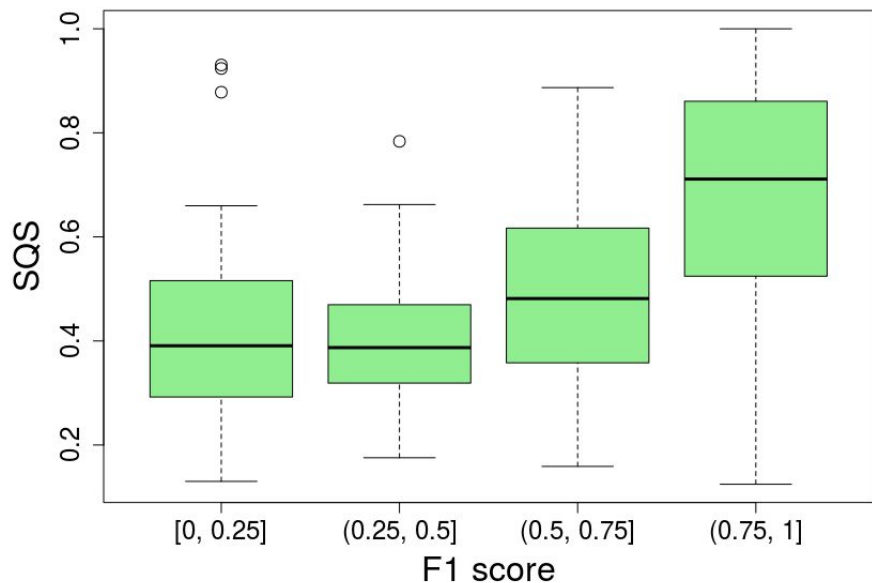
When the crowd workers agree with each other, they also agree with the expert.

But disagreement can have a good reason!

Anca Dumitrache
@anca_dmtrch

Conclusion

Results:

- 433 sentences from FrameNet annotated with CrowdTruth.
- Crowd performance comparable to experts.
- Disagreement is a good indicator of ambiguity in sentences, frames.

Resources:

- Dataset: https://github.com/CrowdTruth/FrameDisambiguation
- CrowdTruth metrics: https://github.com/CrowdTruth/CrowdTruth-core
- CrowdTruth metrics Python package: https://pypi.org/project/CrowdTruth/