

Class 04: SemEval shared tasks

September 7, 2017

History

- Senseval-1
- Senseval-2
- Senseval-3
- SemEval-2007
- SemEval-2010
- SemEval-2012
- SemEval-2013
- SemEval-2014
- SemEval-2015

This year

- Twelve tasks in five main areas
- SemEval-2018 will be the 12th workshop on semantic evaluation and will be located at [TBD].
- Languages represented:
 - English (12), Spanish (XX), Italian (XX), Arabic (XX),

Timeline

25 September	Data ready, evaluation script and baseline system available.
08 January	Evaluation period starts
29 January	Evaluation period ends
26 February	System description paper deadline
2 April	Paper notifications
16 April	Camera-ready submissions due

Approximately 13 weeks to get a system up and running.

Why participate in a shared task ?

- Great learning experience
- All participants get to send a system description paper
- Looks good on CV — both “participated” and “was best in x ”
 - Especially if you’re planning on applying for PhD positions

What next ?

- Describe this year’s tasks
- See if any of them look interesting
- Check out some of the baseline systems

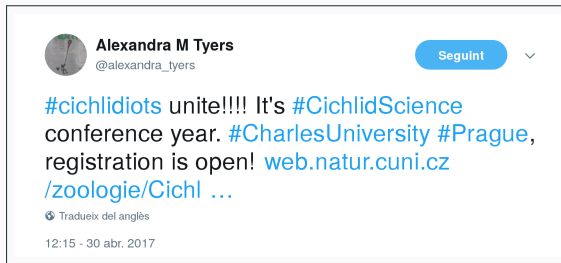
Affect and creative language in tweets

Task 1: Affect in Tweets

Previous datasets / tasks have focussed on simple binary classification of emotion This task looks at how much of each emotion is felt.

For example, given a tweet:

- detect the emotional intensity of {anger, fear, joy, sadness} from 0..1



0.9	Joy
0.0	Fear
0.0	Sadness
0.1	Anger

Languages: English, Arabic, Spanish

Task 2: Multilingual emoji prediction

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
USA	❤️	😍	😂	💕	🔥	😊	😎	✨	💙	😘	📷	🇺🇸	☀️	💜	😏	🏆	😁	🎄	📷	😜
ESP	❤️	😍	😂	💕	😊	😘	💪	😏	👉	🇪🇸	😎	💙	💜	😏	💕	✨	🎵	💕	😁	👆

Task, given a tweet,

- predict which emoji the tweet had beforehand.
- single emoji per tweet.

 **Josep Capdevila#DUI**
@josepcapdevila [Segueix](#) ▾

El País se supera cada día. Ahora el Major de los Mossos se llama Josep Lluís Drapaire
😂😂😂

Traduït del espanyol

EL PAÍS CATALUNYA

En el seu haver-hi, els Mossos expliquen també amb la comunicació. Malgrat la confusió de la primera jornada, la informació sobre els fets va quedar en mans no d'un polític sinó d'un tècnic, el major Josep Lluís Drapaire. El comissari cap ha transmès, coincideixen els experts, un missatge de tranquil·litat. I ha aportat, amb rigor, les dades necessàries per informar a la població, reservant aquells detalls necessaris per esclarir els fets.



 **Josep Capdevila#DUI**
@josepcapdevila [Segueix](#) ▾

El País se supera cada día. Ahora el Major de los Mossos se llama Josep Lluís Drapaire

Traduït del espanyol

EL PAÍS CATALUNYA

En el seu haver-hi, els Mossos expliquen també amb la comunicació. Malgrat la confusió de la primera jornada, la informació sobre els fets va quedar en mans no d'un polític sinó d'un tècnic, el major Josep Lluís Drapaire. El comissari cap ha transmès, coincideixen els experts, un missatge de tranquil·litat. I ha aportat, amb rigor, les dades necessàries per informar a la població, reservant aquells detalls necessaris per esclarir els fets.

Languages: US English, Spanish

Task 3: Irony detection in English tweets

Irony: A trope or figurative language use whose actual meaning differs from what is literally enunciated.

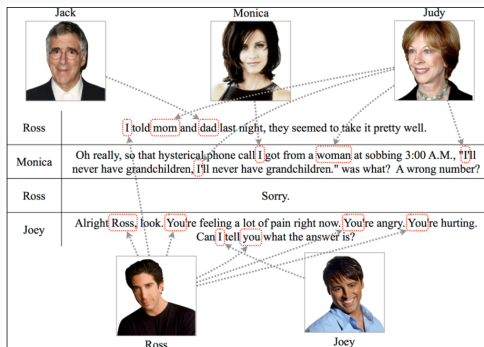
Two tasks:

- Binary classification ironic/not ironic
- Four-way classification:
 - Polarity contrast: I love waking up with migraines #not :'(
 - Situational: Event technology session is having Internet problems. #irony
 - Other: @someuser Yeah keeping cricket clean, that's what he wants #Sarcasm
 - Non-ironic: ...And it's payday. THIS IS A GOOD FRIDAY

Languages: English

Coreference

Task 4: Character identification on multiparty dialogues



									Speaker	Entity ID
0	0	He	PRP	(TOP(S(NP*	he	-	-	Monica_Geller	*	(284)
0	1	's	VBZ	(VP*	be	-	-	Monica_Geller	*	-
0	2	just	RB	(ADVP*)	just	-	-	Monica_Geller	*	-
0	3	some	DT	(NP(NP*	some	-	-	Monica_Geller	*	-
0	4	guy	NN	*)	guy	-	-	Monica_Geller	*	(284)
0	5	I	PRP	(SBAR(S(NP*	I	-	-	Monica_Geller	*	(248)
0	6	work	VBP	(VP*	work	-	-	Monica_Geller	*	-
0	7	with	IN	(PP*))))))	with	-	-	Monica_Geller	*	-
0	8	!	.	*)	!	-	-	Monica_Geller	*	-

Task 5: Counting events and participants in the long tail

Information extraction

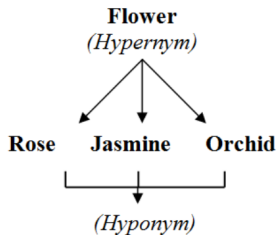
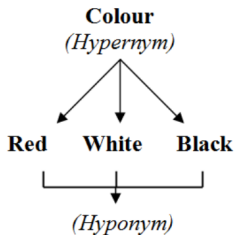
Task 6: Parsing time normalisations

Task 7: Semantic relation extraction and classification in scientific papers

Task 8: Semantic extraction from cybersecurity reports (SecureNLP)

Lexical semantics

Task 9: Hypernym discovery



	Term	Hypernym(s)	Source
English (general)	dog	canine, mammal, animal	WordNet
Spanish (general)	guacamole	salsa para mojar, alimento, salsa	Wikidata
Italian (general)	Nina Simone	musicista, pianista, persona	MultiWiBi

Languages: English, Spanish, Italian

Task 10: Capturing discriminative attributes

- Current models are great at detecting similarity
- But do not predict semantic differences
- If you only know what something is like but can't tell the difference it is a problem for your model



The task:

- Given a triple (cappucino, americano, milk), return
 - 1 is a semantic difference
 - 0 is not a semantic difference

Languages: English

Reading comprehension and reasoning

Task 11: Machine comprehension using commonsense knowledge

My garden was looking a little empty, so I decided I would plant something. I went out and bought tree seeds. I found a spot that looked like it would get enough sunshine. There, I dug a hole for the seeds. Once that was done, I took my watering can and watered the seeds.

A. Why was the tree planted?

- to get enough sunshine
- the garden looks empty
- they were forced
- the soil was moist

B. What was used to dig the hole?

- the wind
- a tablespoon
- a trowel [?]
- their bare hands

Languages: English

Task 11: Machine comprehension using commonsense knowledge

My garden was looking a little empty, so I decided I would plant something. I went out and bought tree seeds. I found a spot that looked like it would get enough sunshine. There, I dug a hole for the seeds. Once that was done, I took my watering can and watered the seeds.

A. Why was the tree planted?

- to get enough sunshine
- the garden looks empty
- they were forced
- the soil was moist

B. What was used to dig the hole?

- the wind
- a tablespoon
- a trowel [?]
- their bare hands

Languages: English

Task 11: Machine comprehension using commonsense knowledge

My garden was looking a little empty, so I decided I would plant something. I went out and bought tree seeds. I found a spot that looked like it would get enough sunshine. There, I dug a hole for the seeds. Once that was done, I took my watering can and watered the seeds.

A. Why was the tree planted?

- to get enough sunshine
- the garden looks empty
- they were forced
- the soil was moist

B. What was used to dig the hole?

- the wind
- a tablespoon
- a trowel [?]
- their bare hands

Languages: English

Task 12: Argument reasoning comprehension task

Given an argument, with a reason and a claim, the goal is to choose the correct implicit reasoning from two options.

Languages: English

Practical

<http://alt.qcri.org/semeval2018/index.php?id=tasks>

- Choose a task
- Read the description
- Look at the provided data
- Run the baseline system
- Think of ways to improve the baseline