Results of the PolEval 2019 Shared Task 1
Recognition and Normalization of Temporal Expressions

Jan Kocoń, Marcin Oleksy, Tomasz Bernaś,
Michał Marcińczuk

Department of Computational Intelligence
Faculty of Computer Science and Management
Wroclaw University of Science and Techology
jan.kocon@pwr.edu.pl

May 31, 2019

Politechnika
Wrocławska

Goal fo the task

- Recognize temporal expressions (timexes) in a text.
- Normalize the expressions — generate relative (local) and absolute (global) values.

- **When** something happens?
- **How long** something lasts?
- **How often** something occurs?
- 4 classes: date, time, duration, set.

paru godzin z siostrzenicą Dębickiego Madzia nie miała zajęcia.

Każdego poranku trapiła ją chęć wyjścia, ale - po co i dokąd? Więc siedziała samotna w domu trwożąc się, że nic nie robi, i czekając na list od Zdzisława.

" Dziś z pewnością przyjdzie - myślała. - Nie było z rana , więc będzie po południu ... . Nie było dziś , więc jutro ... ."

Figure 1: Examples of temporal expressions taken from Polish Corpus of Wroclaw University of Technology (KPWr) [1]

[1] Bartosz Broda et al. "KPWr: Towards a Free Corpus of Polish". In: *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. Ed. by Calzolari Nicoletta et al. Istanbul, Turkey: European Language Resources Association (ELRA), May 25, 23. ISBN: 978-2-9517408-7-7

TimeML

- Specification language for annotation of events and temporal expressions:
  - temporal expressions (anchors for events),
  - event mentions,
  - relations between events,
- Accepted as ISO standard [2]
- TimeML is a widely used standard:
  - Events – 6 languages,
  - Temporal expressions – 13 languages.

---

[2] James Pustejovsky et al. "ISO-TimeML: An International Standard for Semantic Annotation". In: *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Valletta, Malta: European Language Resources Association (ELRA), May 2010. ISBN: 2-9517408-6-7

Normalisation of temporal expressions

- Transformation into a form *understandable* by a computer
- Two stages:
  - Local Value (LVAL) — expression only (LTIMEX [3])
  - Global Value (VAL) — reasoning using the entire document (TimeML, TIMEX3 [4])
- Previous works: adaptation of normalisation guidelines to Polish [5] and HeidelTime-like method of normalising temporal expressions [6].

_____

[3] Paweł Mazur. "Broad-Coverage Rule-Based Processing of Temporal Expressions". rozprawa doktorska. Politechnika Wrocławska, 2012

[4] Roser Saurí et al. *TimeML Annotation Guidelines, Version 1.2.1.* 2006

[5] Jan Kocoń et al. "Temporal Expressions in Polish Corpus KPWr". In: *Cognitive Studies — Etudes Cognitives* 15 (2015)

[6] Jan Kocoń and Michał Marcińczuk. "Supervised approach to recognise Polish temporal expressions and rule-based interpretation of timexes". In: *Natural Language Engineering* 23.3 (2017), 385–418. DOI: 10.1017/S1351324916000255

Normalisation of temporal expressions (examples)

timex (PL)	timex (EN)	LVAL	VAL
1995-06-06	1995-06-06	1995-06-06	1995-06-06
6 czerwca	6th June	xxxx-06-06	1995-06-06
dwa dni temu	two days ago	-0000-00-02	1995-06-04
dwa tygodnie temu	two weeks ago	-0000-W02	1995-05-23
sześćdziesiątym ósmym	sixty-eight	xx68	1968
8 wieczorem w piątek	8pm on Friday	xxxx-Wxx-5T20:00	1995-05-26T20:00
o 8 w piątek	at 8 on Friday	xxxx-Wxx-5t08:00	1995-05-26t08:00
następna środa	next Wednesday	>D3	1995-05-24
dziewięć miesięcy	nine months	P9M	P9M

Table 1: LVAL and VAL for example temporal expressions. Reference date for derermination of VAL is 1995-06-06.

Language	SemEval2013 EN and SP		KPWr PL	
Documents and tokens				
Documents	2 907		1 635	
Tokens	835 885		447 576	
Tok./doc.	288		273	
Annotations				
date	13 852	(74.09%)	4 391	(71.80%)
time	2 817	(15.07%)	928	(15.17%)
duration	1 888	(10.10%)	653	(10.68%)
set	139	(0.74%)	144	(2.35%)
Total	18 696		**6 116**	

Table 2: Number of annotations in KPWr vs corpora shared within SemEval2013.

Data format

```
<DOCID>344245.xml</DOCID>
<DCT>
<TIMEX3 tid="t0" functionInDocument="CREATION_TIME"
type="DATE" value="2006-12-16">
</TIMEX3>
</DCT>
<TEXT>
<TIMEX3 tid="t1" type="DATE" value="2006-12-16">Dziś</TIMEX3> Creative
Commons obchodzi czwarte urodziny - przedsięwzięcie ruszyło dokładnie
<TIMEX3 tid="t2" type="DATE" value="2002-12-16">16 grudnia 2002</TIMEX3>
w San Francisco. (...) Z kolei w <TIMEX3 tid="t4" type="DATE"
value="2006-12-18">poniedziałek</TIMEX3> ogłoszone zostaną wyniki
głosowanie na najlepsze blogi. W ciągu <TIMEX3 tid="t5" type="DURATION"
value="P8D">8 dni</TIMEX3> internauci oddali ponad pół miliona głosów.
Z najnowszego raportu Gartnera wynika, że w <TIMEX3 tid="t6" type="DATE"
value="2007">przyszłym roku</TIMEX3> blogosfera rozrośnie się
do rekordowego rozmiaru 100 milionów blogów. (...)
</TEXT>
```

Testing dataset

Set of **100 documents** in the following format:

```
<DOCID>344245.xml</DOCID>
<DCT><TIMEX3 tid="t0" functionInDocument="CREATION_TIME" type="DATE"
value="2006-12-16"></TIMEX3></DCT>
<TEXT>
Dziś Creative Commons obchodzi czwarte urodziny - przedsięwzięcie ruszyło
dokładnie 16 grudnia 2002 w San Francisco. (...) Z kolei w poniedziałek
ogłoszone zostaną wyniki głosowanie na najlepsze blogi. W ciągu 8 dni
internauci oddali ponad pół miliona głosów. Z najnowszego raportu Gartnera
wynika, że w przyszłym roku blogosfera rozrośnie się do rekordowego
rozmiaru 100 milionów blogów. (...)
</TEXT>
```

Normalisation of timexes

- State-of-the-art: rule-based methods.
- Most approaches: complex and specific rules.
- Best SemEval2013 system: HeidelTime – 326 rules, VAL F1: 77,61% [7].
- Adaptation to Polish in Liner2 system: 224 rules, VAL F1: 66,71% [6].
- Liner2 improved with Cascade of Partial Rules: 167 rules, VAL F1: 77,23% [8]

---

[7] Naushad UzZaman et al. "Semeval-2013 task 1: Tempeval-3: Evaluating time expressions, events, and temporal relations". In: *Second Joint Conference on Lexical and Computational Semantics (* SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013)*. Vol. 2. 2013, pp. 1–9

[6] Jan Kocoń and Michał Marcińczuk. "Supervised approach to recognise Polish temporal expressions and rule-based interpretation of timexes". In: *Natural Language Engineering* 23.3 (2017), 385–418. DOI: 10.1017/S1351324916000255

[8] Jan Kocoń and Michał Marcińczuk. "Improved Recognition and Normalisation of Polish Temporal Expressions". In: *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*. 2017, pp. 387–393

- How many entities are correctly identified
  - $P = \frac{|Sys_{entity} \cap Ref_{entity}|}{Sys_{entity}}$
  - $R = \frac{|Sys_{entity} \cap Ref_{entity}|}{Ref_{entity}}$
- If the extents for entities are correctly identified:
  - Strict match: |16 January 2019| vs |16 January 2019|
  - Relaxed match: |January 2019| vs |16 January 2019|
- (if relaxed match) How many attributes are correctly identified:
  - $attrP = \frac{|\forall x|x \in (Sys_{entity} \cap Ref_{entity}) \wedge Sys_{attr}(x)==Ref_{attr}(x)|}{Sys_{entity}}$
  - $attrR = \frac{|\forall x|x \in (Sys_{entity} \cap Ref_{entity}) \wedge Sys_{attr}(x)==Ref_{attr}(x)|}{Ref_{entity}}$
  - $attrF = \frac{2PR}{P+R}$

Participating systems

- Alium – prepared by Marcin Bodziak
    - rule-based recognition and normalisation
    - complex rules incorporating words, lemmas, PoS, word masks (digits, special characters)
    - 420 rules in a 6-level hierarchy
- Liner2 – baseline [8]
    - did not take part in the competition
    - recognition based on machine learning (CRF)
    - rule-based normalisation (CPR)
    - 167 rules in cascade

[8] Jan Kocoń and Michał Marcińczuk. "Improved Recognition and Normalisation of Polish Temporal Expressions". In: *Proceedings of the International Conference Recent Advances in Natural Language Processing, RANLP 2017*. 2017, pp. 387–393

Results

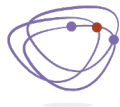| Strict Match	F1	P	R
Alium	58.81	58.91	58.72
Liner2-baseline	87.63	86.17	89.14
Relaxed Match	F1	P	R
Alium	86.49	86.63	86.35
Liner2-baseline	91.19	89.67	92.76
Attribute F1	Value	Type	
Alium	68.70	80.23	
Liner2-baseline	76.96	87.79	

Table 3: Results of the recognition and normalization of temporal expressions obtained by Alium system, compared to results obtained by Liner2 system.