Masteratul de Lingvistică Computațională

Curs: *Introducere in Lingvistica Computațională*

Curs 9

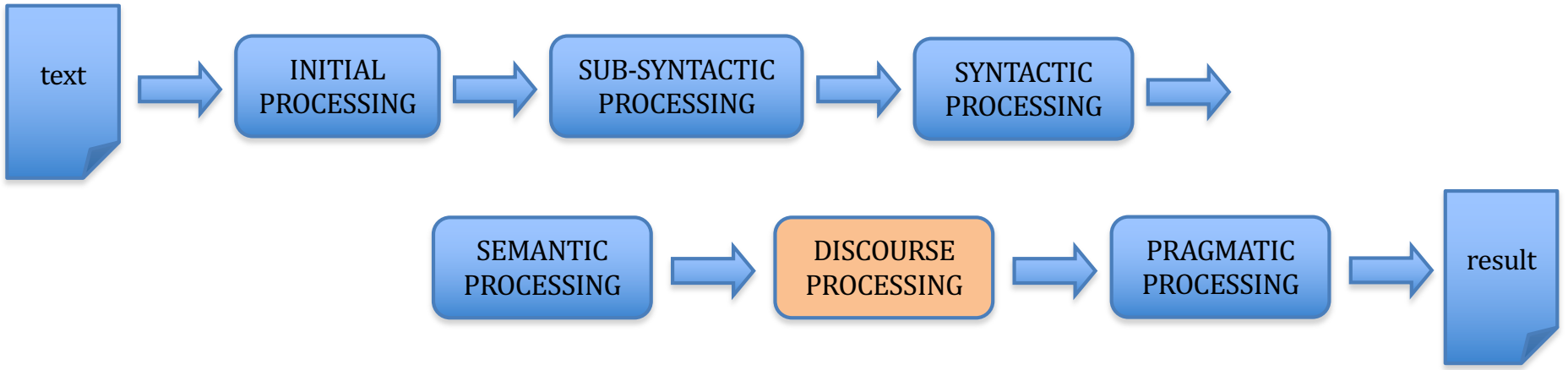
Time in texts:

Temporal Logic, TimeML, Time Yards

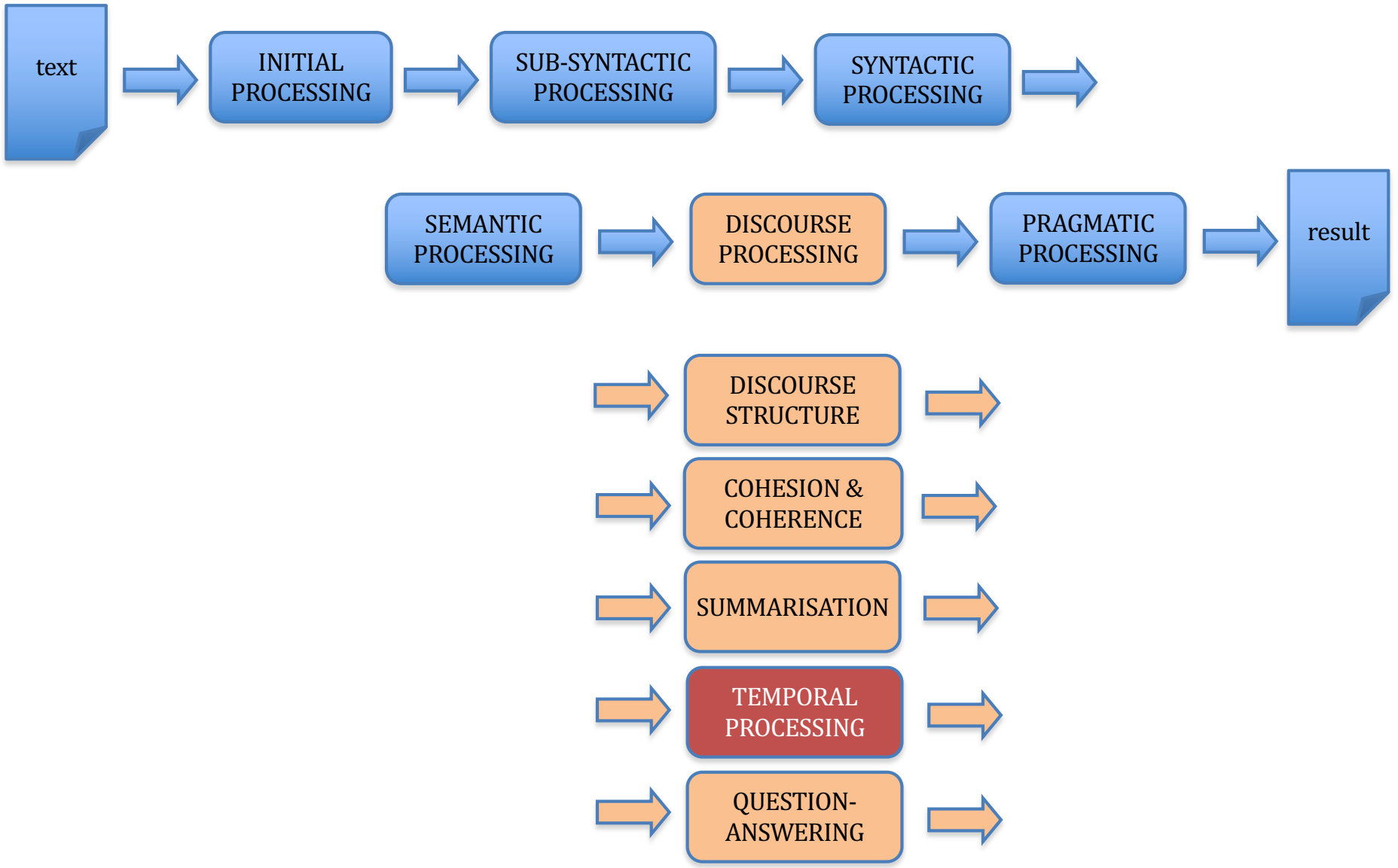
Curs: Dan Cristea

Seminarii & proiect: Mihaela Onofrei, Dan Cristea

The discourse layer



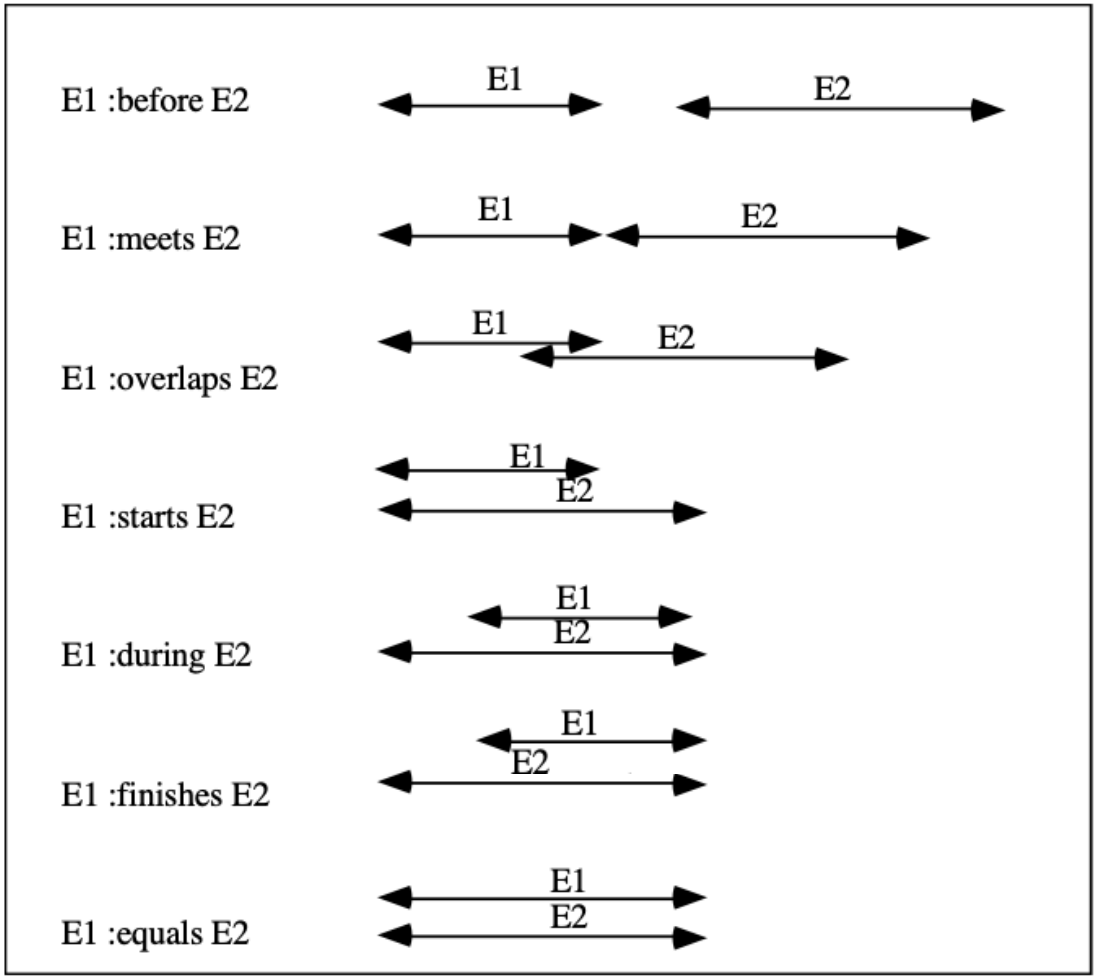
Temporal processing



James Allen’s Temporal Reasoning

Allen's theory on time and events

- O logică a intervalelor temporale



Allen's theory on time and events

- A logic of time => simbolic reasoning about time, as it is mentioned in texts

$$\text{BEFORE}(t1, t2) \,\&\, \text{BEFORE}(t2, t3) \Rightarrow \text{BEFORE}(t1, t3), \qquad \text{(T.1)}$$

$$\text{MEETS}(t1, t2) \,\&\, \text{DURING}(t2, t3) \Rightarrow \qquad \text{(T.2)}$$
$$(\text{OVERLAPS}(t1, t3) \vee \text{DURING}(t1, t3) \vee \text{MEETS}(t1, t3)).$$

Allen, J. (1984). Towards a General Theory of Action and Time, Artificial Intelligence 23: 123–54.
http://www.cs.ucf.edu/~lboloni/Teaching/EEL6938_2007/papers/Allen-GeneralTheoryActionTime.pdf

The TimeML model

TimeML

- An approach to temporal annotation in texts
  - EVENT <= marks mentions of events
    - typically associated to verbs
    - but nouns may also denote events: crash (*killed by the crash*), bankraft (*the last year bankraft of the company*), marriage (*I was invited to their marriage*)…

James Pustejovsky *et al.* (2003). TimeML: Robust Specification of Event and Temporal Expressions in Text, AAAI Technical Report SS-03-07

Sauri, R., Littman, J., Knippen, B., Gaizauskas, R., Setzer, A., and Pustejovsky, J. (2006). TimeML Annotation Guidelines Version 1.2.1. https://catalog.ldc.upenn.edu/docs/LDC2006T08/timeml_annguide_1.2.1.pdf

TimeML: EVENT and MAKEINSTANCE

EVENT marks up events.
MAKEINSTANCE indicates different instances of a given event

- should have bought

```
should have
<EVENT eid="e1" class="OCCURRENCE">bought</EVENT>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PAST"
aspect="PERFECTIVE" modality="SHOULD"/>
```

- did not teach

```
did not
<EVENT eid="e1" class="OCCURRENCE">teach</EVENT>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PRESENT"
aspect="NONE" polarity="NEG"/>
```

- must not teach twice

```
must not
<EVENT eid="e1" class="OCCURRENCE">teach</EVENT>
<SIGNAL sid="s1">twice</SIGNAL>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PRESENT"
aspect="NONE" polarity="NEG" modality="MUST" signalID="s1"
cardinality="2"/>
```

TimeML: TIMEX3

Marks up explicit temporal expressions, such as times, dates, durations, etc.

- no more than 60 days

```
<TIMEX3 tid="t1" type="DURATION" value="P60D"
mod="EQUAL_OR_LESS">no more than 60 days</TIMEX3>
```

- the dawn of 2000

```
<TIMEX3 tid="t2" type="DATE" value="2000" mod="START">the dawn
of 2000</TIMEX3>
```

- twice a month

```
<TIMEX3 tid="t3" type="SET" value="P1M" freq="2X">twice a
month</TIMEX3>
```

- daily

```
<TIMEX3 tid="t5" type="SET" value="P1D
quant="EVERY">daily</TIMEX3>
```

TimeML: SIGNAL

Annotate function words that indicate how temporal objects are related to each other.

- must not teach twice

```
must not
<EVENT eid="e1" class="OCCURRENCE">teach</EVENT>
<SIGNAL sid="s1">twice</SIGNAL>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PRESENT"
aspect="NONE" polarity=”NEG" modality="MUST" signalID="s1"
cardinality="2"/>
```

- two weeks from June 7, 2003

```
<TIMEX3 tid="t6" type="DURATION" value="P2W" beginPoint="t61"
endPoint="t62">two weeks</TIMEX3>
<SIGNAL sid="s1">from</SIGNAL>
<TIMEX3 tid="t61" type="DATE" value="2003-06-07">June 7,
2003</TIMEX3>
<TIMEX3 tid="t62" type="DATE" value="2003-06-21"
temporalFunction="true" anchorTimeID="t6”/>
```

TimeML: TLINK

Temporal link: represents the relation between two temporal elements.

- John taught twice on Monday but only once on Tuesday.

```
John
<EVENT eid="e1" class="OCCURRENCE">taught</EVENT>
<SIGNAL sid="s1">twice</SIGNAL>
<SIGNAL sid="s2">on</SIGNAL>
<TIMEX3 tid="t1" type="DATE" value="xxxx-wxx1">Monday</TIMEX3>
but only
<SIGNAL sid="s3">once</SIGNAL>
<SIGNAL sid="s4">on</SIGNAL>
<TIMEX3 tid="t2" type="DATE" value="xxxx-wxx2">Tuesday</TIMEX3>
<MAKEINSTANCE eiid="ei1" eventID="e1" tense="PAST"
aspect="NONE" signalID="s1" cardinality="2"/>
<MAKEINSTANCE eiid="ei2" eventID="e1" tense="PAST"
aspect="NONE" signalID="s3" cardinality="1"/>
<TLINK eventInstanceID="ei1" signalID="s2" relatedToTime="t1"
relType="IS_INCLUDED"/>
<TLINK eventInstanceID="ei2" signalID="s4" relatedToTime="t2"
relType="IS_INCLUDED"/>
```

Related work

- TARSQI: a toolkit that recognises:
  - temporal expressions (TimeEx)
  - events
  - relations between them

Verhagen & Pustejovsky: Temporal Processing with the TARSQI Toolkit, Coling 2008

Related work: Narrative Containers

- Assumed temporal windows that provide left and right boundaries for when unanchored events most likely occurred; size parameterized by Document Creation Time (DCT) and the periodicity of the source
- Narrative Scope: the timespan between the earliest and latest events described in the document
- Narrative Time: the current temporal anchor for events in a document; temporal focus shifts as the reader moves through the document

=> Shows how TimeML can reflect the manner in which events are linked (related) to the DCT

Related work: Narrative Containers

- The relations of the model of wrt to:
  - Channel: is the document written or spoken?
  - Production circumstances: how was the document distributed? broadcast, newswire, daily publication
  - Style: what format was used to present the information?
  - Presence of a temporal anchor: whether an article contained a Narrative Time in the first sentence of the document

Pustejovsky&Stubbs (2011). Increasing Informativeness in Temporal Annotation, *Proceedings of the Fifth Law Workshop (LAW V)*, Portland

However…

- Fiction barely mentioned in these studies
  - most interest attached to process newswires and broadcasted information
- Less or no attention given to the link between events and characters involved
  - and there is more than the mere role of agent in an event:

    John met Ana when exiting the Opera House. He said: Amazing voices!
  - here Ana was the person that heard John's exclamation, a fact which is not explicitly signaled

The Time Yards Model

Motivation and goals

- Large texts are different then news
  - news: sequences of events, opinions, commentaries
  - novels: events are seldom presented in non-sequential order, same events are interpreted from different perspectives
- Characters participate in events
- Each has her/his own trajectory/destiny/evolution
- Although characters' trajectories intersect in time, one would like to distinguish/separate them

Motivation and goals

- A computational representation of temporal phenomena occurring in literary texts
- Prepare investigation tools that would make possible to tackle these phenomena by automatic analysis
- Make explicit things that readers perceive in texts:
  - changes of the current direction of time,
  - flashbacks and flashforwards,
  - rupture points,
  - perspectives of different narrators on the same story

We are interested in connected sequences of events

- Other approaches in NLP
  - *TimeLine*: a representation of events which are chronologically ordered; in general, specific for an entity (a character or participant in an action, a geographical place or region)

Chambers and Jurafsky (2008). Unsupervised Learning of Narrative Event Chains. In ACL.

  - *StoryLine*: groups of interacting TimeLines, or mergers of two or more TimeLines where the same characters or entities are taking part in the action. A kind of text structure which is not related to the flow of time.

Laparra, Aldabe and Rigau (2015). From TimeLines to StoryLines: A preliminary proposal for evaluating narratives. In ACL-IJCNLP.

Tash Aw: *Map of the invisible world*

In novels, time does not flow linearly on…

Our perception of time

- *She remembered the words that sometimes, at the time of adolescence, her mother told her. "You are a person apprehended, right?...*

Tash Aw: Map of the invisible world

Our perception of time

Our perception of time

- *In the kitchen, waiting for the water to boil, Margaret wondered how was it that she jumped so naturally to embrace Adam. She could not realize what had made her feel so happy to see him again.*

Tash Aw: Map of the invisible world

Very often, the author carries us backward in time

- [*In the kitchen, waiting for the water to boil, Margaret wondered how was it that*]

Margaret: story time, in the kitchen, wondering smth.

- [*she jumped so naturally to embrace Adam.*]

Margaret & Adam: in the past, somewhere, Margaret embraces Adam

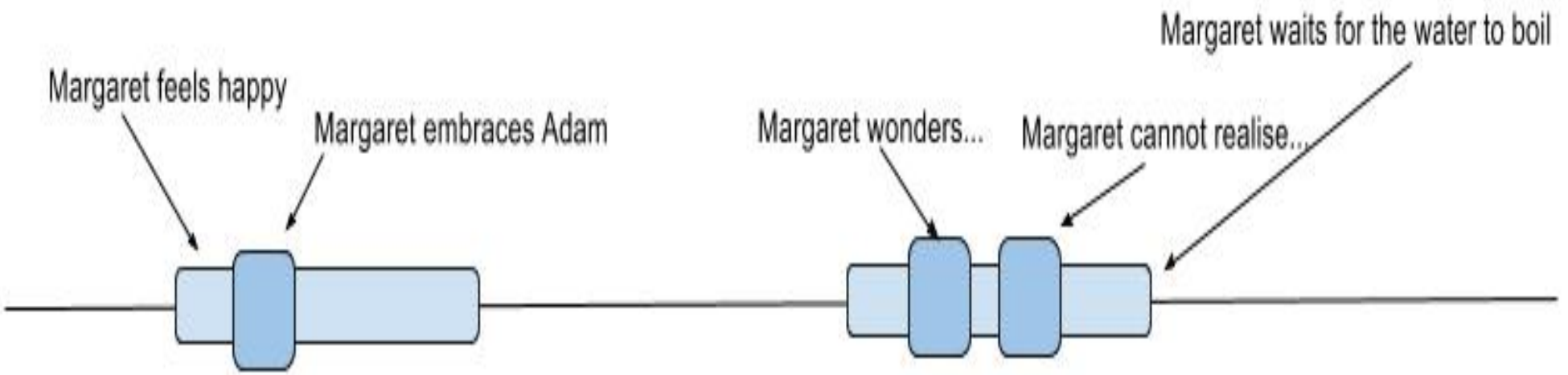
- [*She could not realise what had made her*]

Margaret: story time, in the kitchen, not realizing smth

- [*feel so happy to see him again.*]

Margaret & Adam: in the past, somewhere, Margaret feels happy

Our perception of time



Time Segments

- A TS groups a sequence of events of one character or a stable group of characters over a period of story time, that is uninterruptedly told in a span of text.
  - one TS is necessarily told in a continuous span of text
  - in a TS, the story time evolves linearly
  - the group of characters involved is stable
  - there is only one narrator (relator)

Recognition of TS's border points

- where there is a suddern change in the flow of time

OR

- where there is a change in the group of characters involved

OR

- where the relator changes (change of perspective)

Time Tracks

- A TT displays the story development of one character or a stable group of characters over a period of story time.
  - TTs are deciphered by the reader as having a unity of characters, action and development
  - One TT can be narrated by one or more relators
  - TTs are not necessarily characterised by a unity of place
  - One TT can be composed of TSs which may not be adjacent in the text

Border points of TTs

- Time tracks can be bordered by:
  - *start point* – where the whole story starts,
  - *end point* – where it finishes,
  - *join points* – in which two different time tracks unify,
  - *split points* – where one time track splits into two separate time tracks.

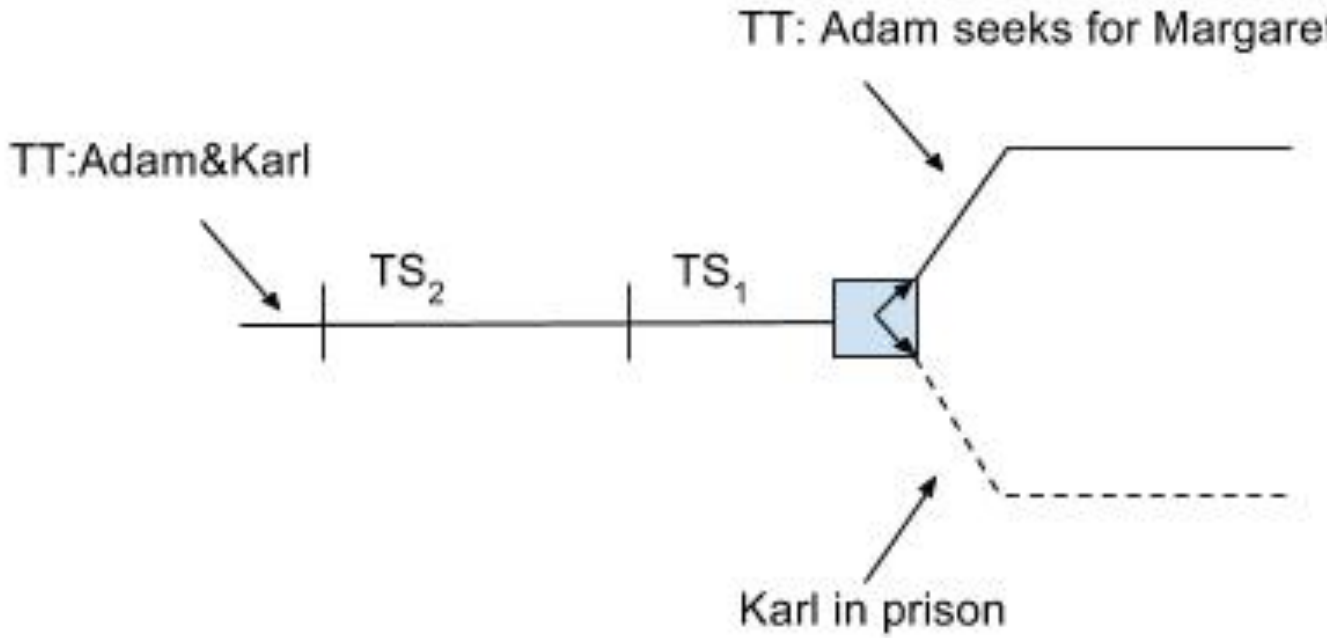
The Time Yards model

- A book is a sequence of Time Segments (TSs)
- TSs make up Time Tracks (TTs)
- A composition of TTs is a Time Yard (TY): a representation (resembling a diagram) of the characters' destinies in a text, unfolded on the story time axis

Dan Cristea, Andreea Macovei (2017). Why Time Resembles Rail Yards? A Way to Look at Time in Text Books, in Romanian Journal of Information Science and Technology, Vol. 19, 1-2, 2016, pp. 31-43

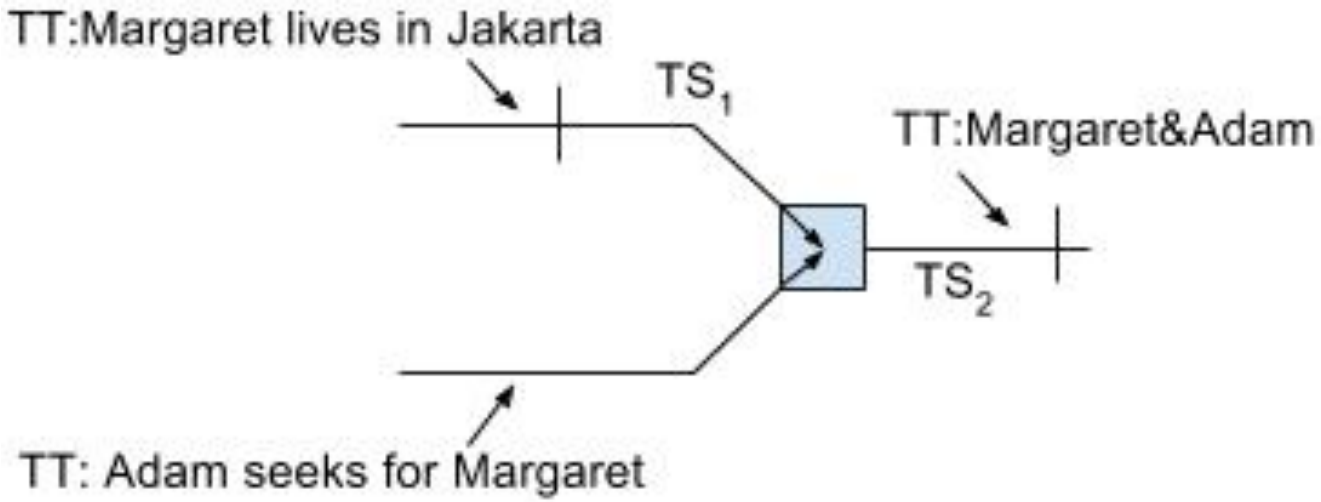
Where 2 TTs split: *Split Points*

- TS1:[Up to climb into the truck and disappear after the tarpaulin fall down,]
- TS2:[he followed them slowly, with his clumsy gait.]

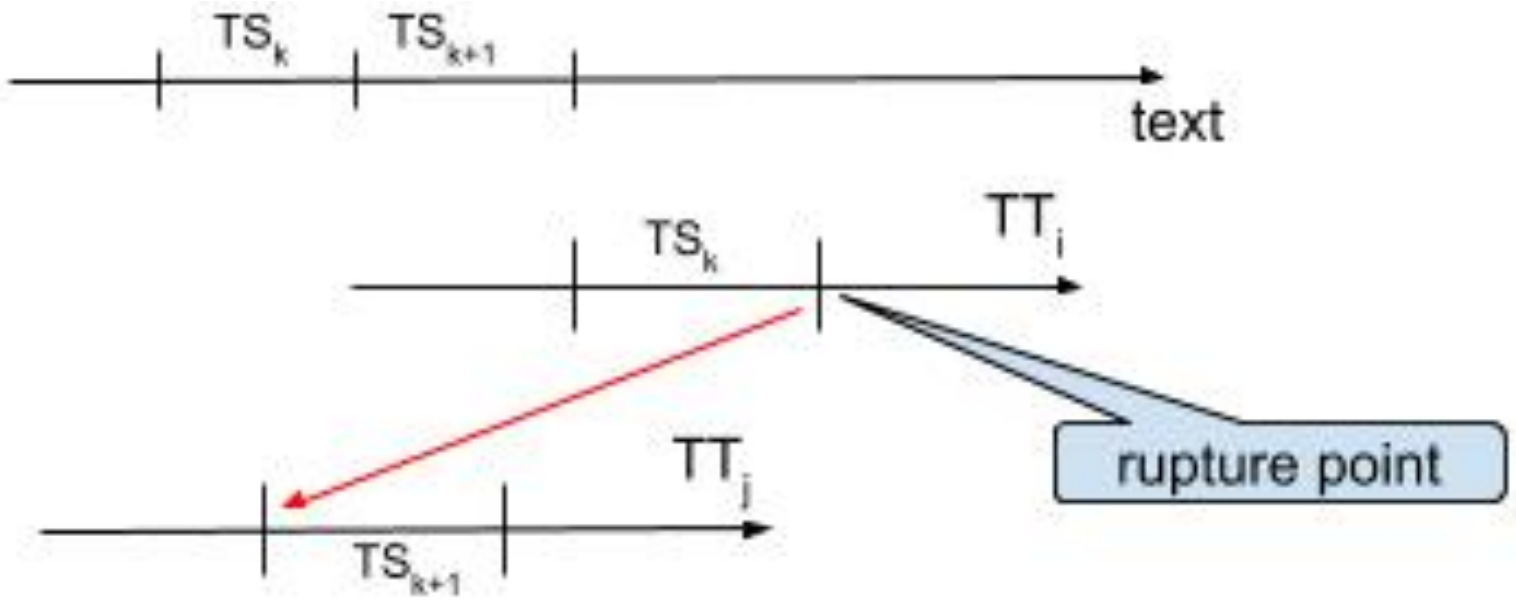


Where 2 TTs join: *Join Points*

- TS1:[It was only when she came to the doorway that she realized]
- TS2:[someone was there, lying on the front steps. He was a boy, a teenager, almost crouched in the womb.]

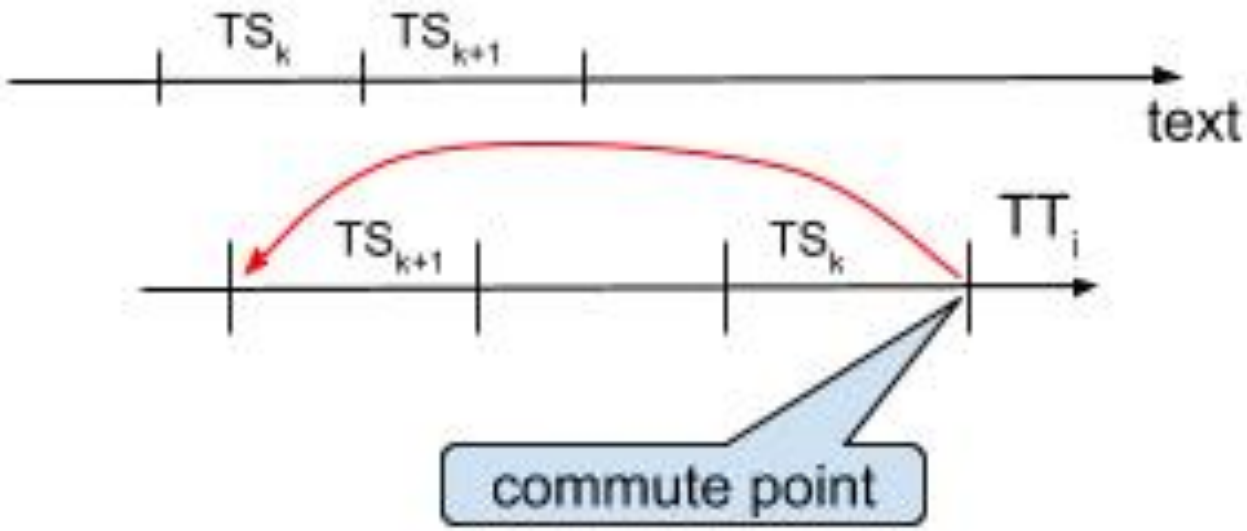


Jumping from one TT onto another: *Rupture Points*



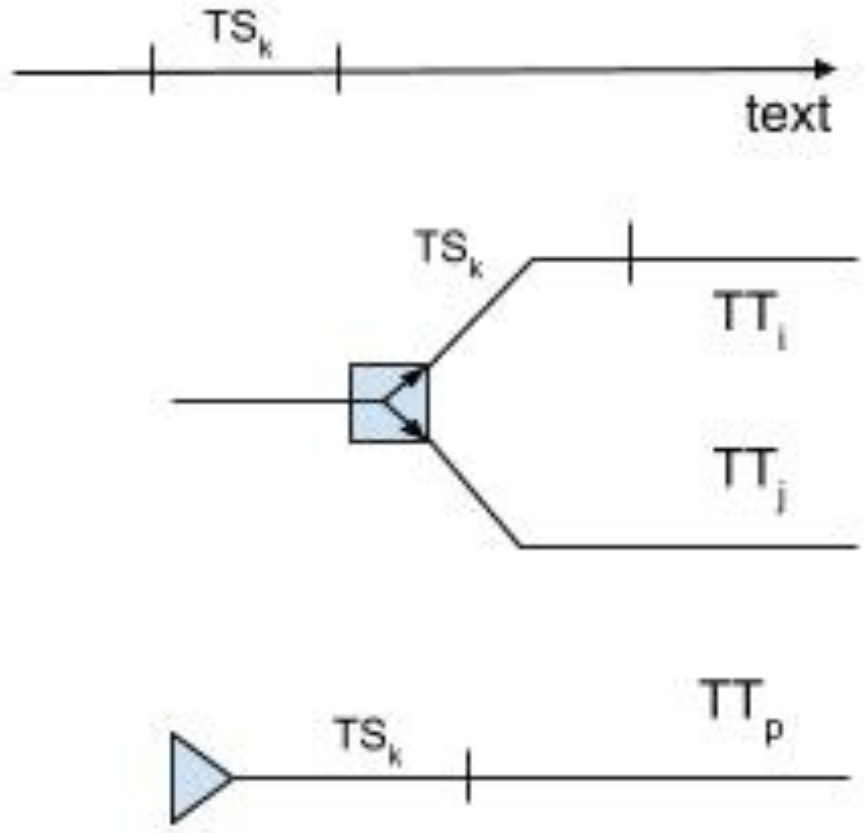
Where the story moves back/forth on the same TT: *Commute Points*

- Or it is said from a different perspective (by a different storyteller)



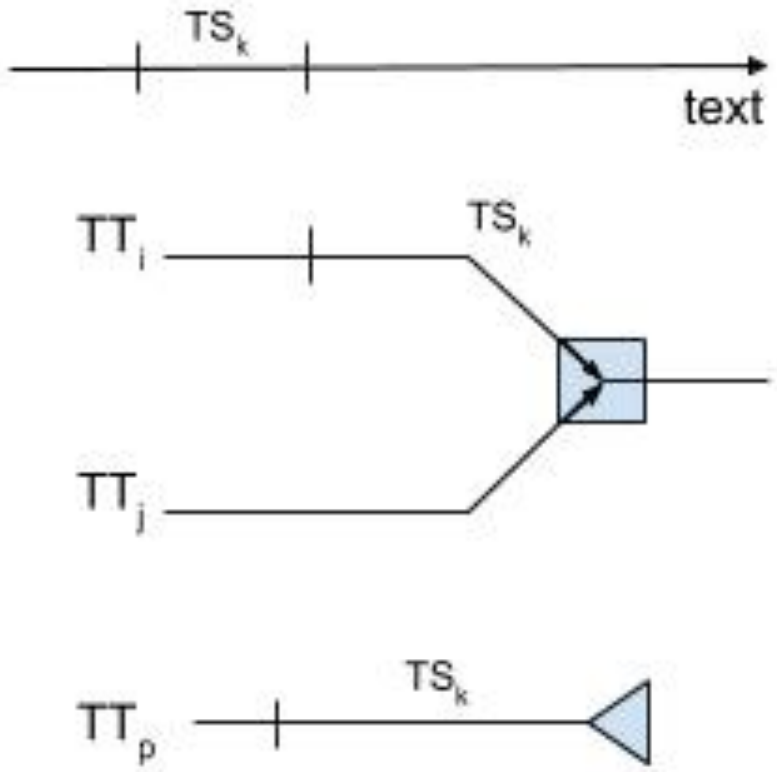
A TT can start in a *Split* or *Start Point*

- On the left side



A TT can end in a *Join* or *End Point*

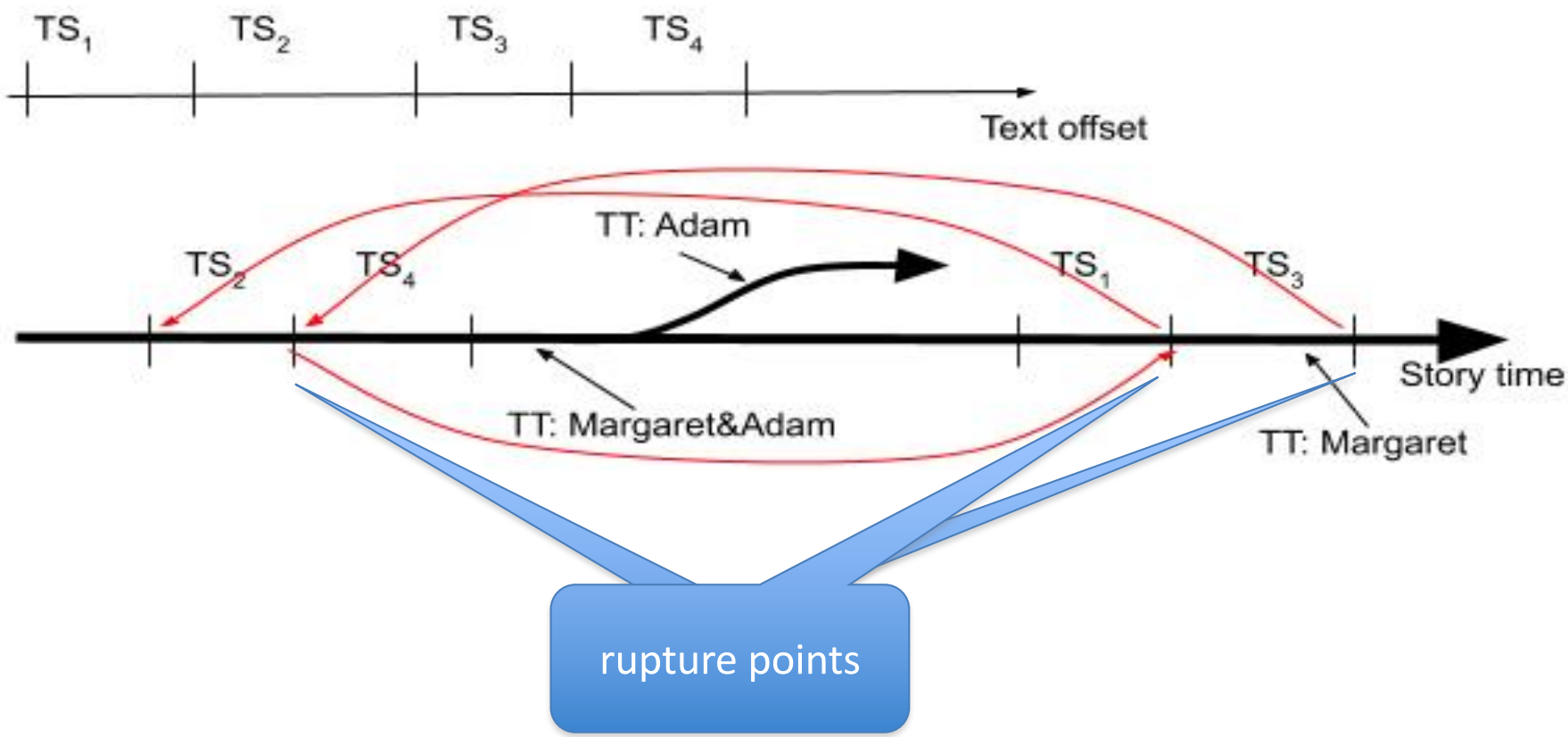
- On the right side.



Margaret in the kitchen…

- TS1:[*In the kitchen, waiting for the water to boil, Margaret wondered how was it that*]
- TS2:[*she jumped so naturally to embrace Adam.*]
- TS3:[*She could not realize what had made her*]
- TS4:[*feel so happy to see him again…*]”

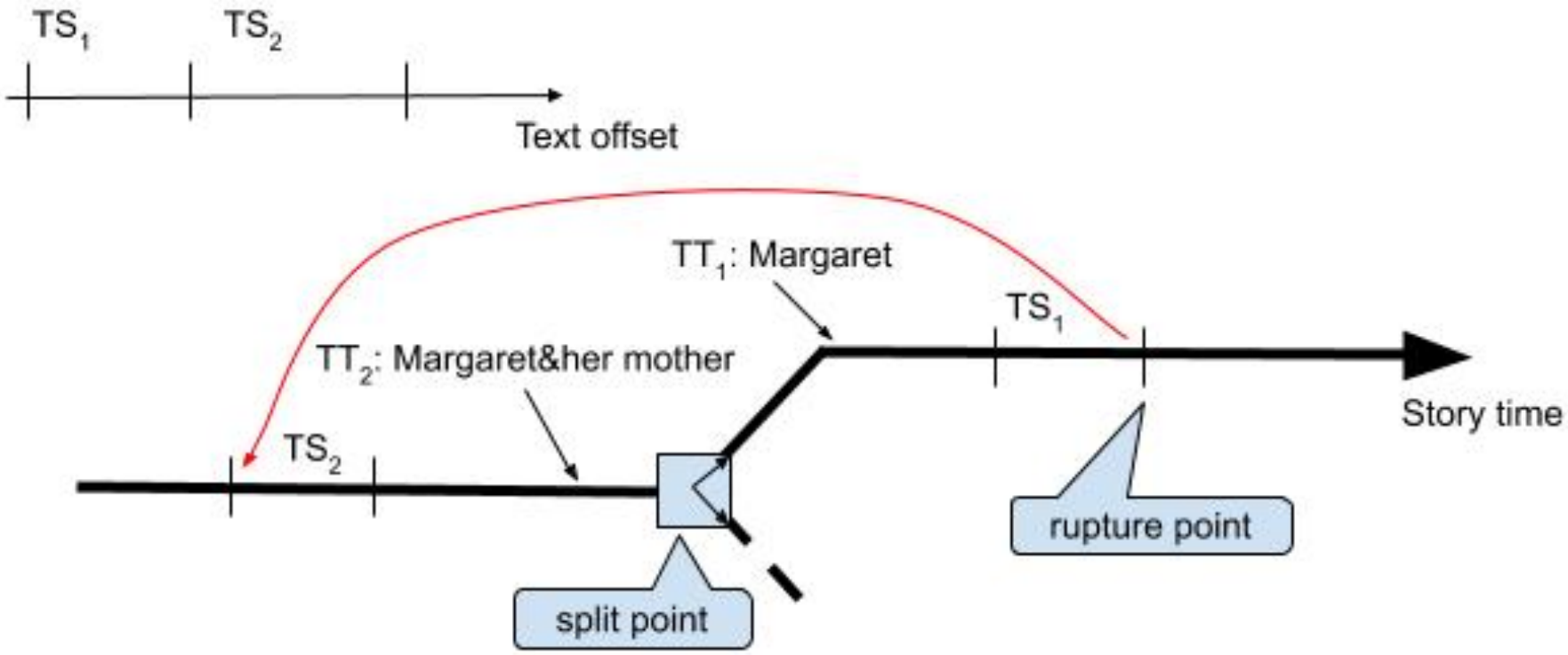
TTs and TSs in Margaret & Adam's episode



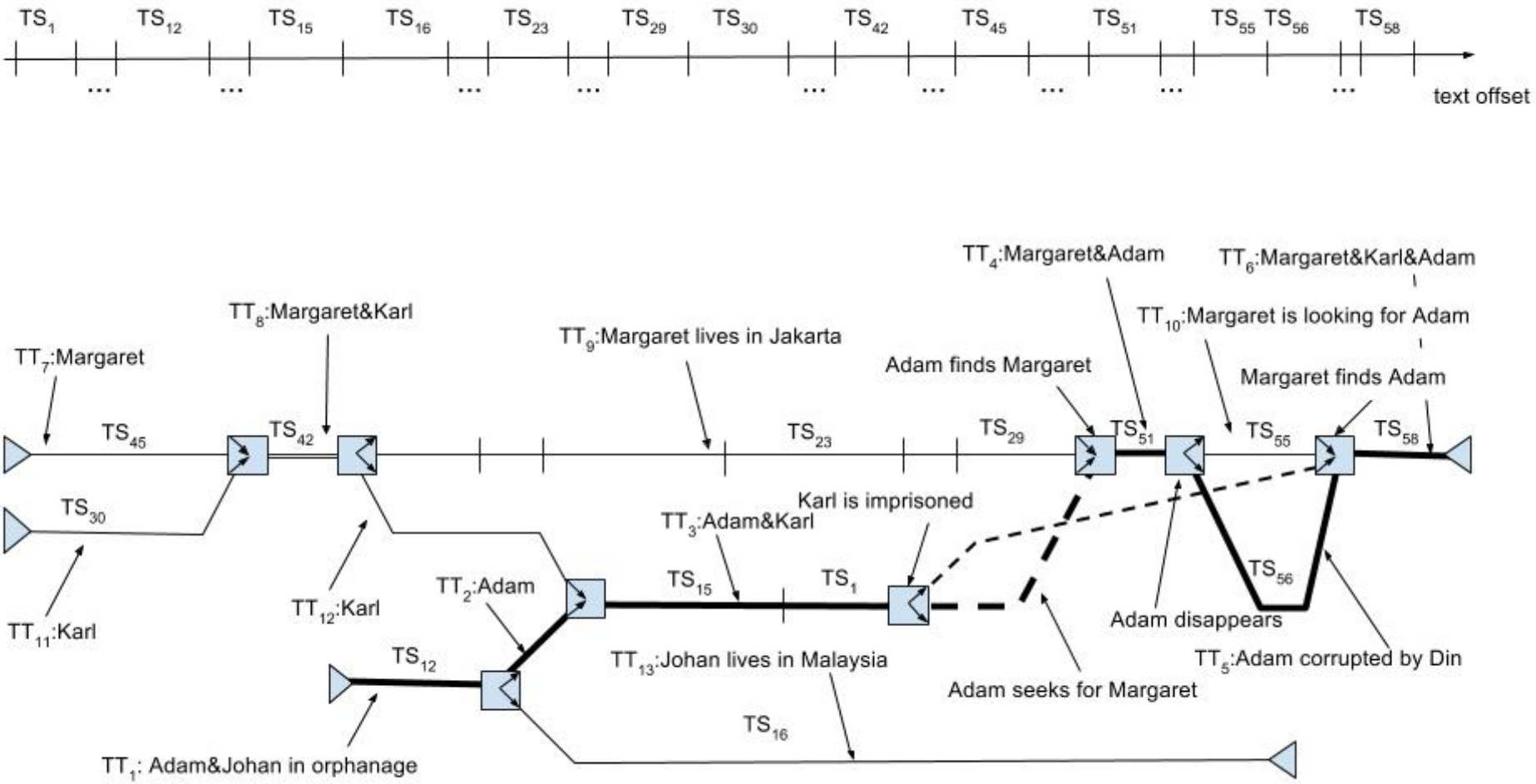
Margaret and her mother…

- TS1:[*She remembered the words*]
- TS2:[*that sometimes, at the time of adolescence, her mother told her. "You are a person apprehended, right?...*]

TTs and TSs in Margaret & her mother's episode



Tash Aw's "Map of the invisible world" bird's eye Time Yard



Reading the TY diagrams

- Tracking characters' destinies (developments)

  => in bold in the figure: Adam's trajectory

  => a mere selection operation in the set of TT elements

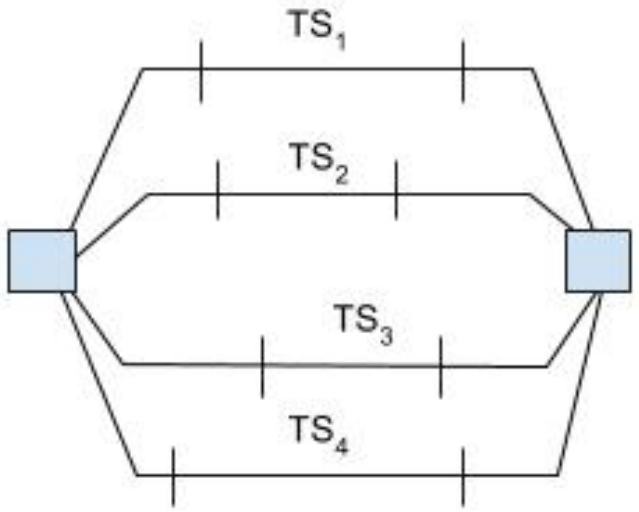
  => produce focused summaries

Not shown in the TY diagram

- Ambiguous temporal placements: when the exact order of TSs cannot be determined

  => more TSs belong to the same TT

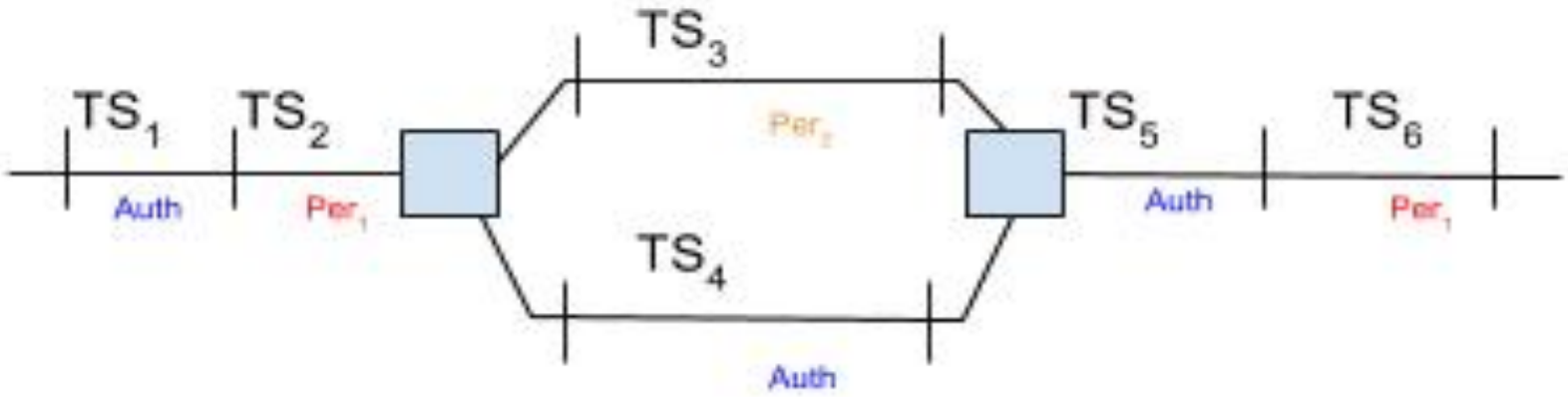
  => but their order is undefined



Not shown in TY diagrams

- Personal interpretations:

  => different episodes (TSs) are told by different relators



An XML notation –

complementing those of TimeML

- Elements marking Time Tracks:

  `<TT NAME ACTORS ID LEFT-EP RIGHT-EP/>`

  –LEFT-EP: “START” | ID of a SPLIT end-point

  –RIGHT-EP: “STOP” | ID a JOIN end-point

- Elements marking Time Segments

  `<TS ID IN-TT TYPE PER>...</TS>`

  –IN-TT: ID of a TT

  –TIME: ID of a TIMEX3 element

  –TYPE: NARration, REMembers, SUPposition, GENeral knowledge, FICtions

  –PERspective: a character (entity)

An XML notation

- Elements marking time constraints

  `<TIMEX ID FROM REL TO/>`

  –FROM and TO: IDs of TS

  –REL: BEFORE, IMMEDIATELY-BEFORE, AFTER, IMMEDIATELY-AFTER, SIMULTANEOUSLY, INCLUDES

- Elements marking endpoints

  `<ENDPOINT ID TYPE TT1 TT2/>`

  –TYPE: “JOIN” or “SPLIT”

  –TT1 and TT2: IDs of the two TTs that either join or split

The corpus

- Tash Aw: *Map of the invisible world* (Romanian version) to be used in training a TY parser
- A few pages were annotated by a class of master students in CL
  - interest in evidencing inter-annotator agreement
- A research still under development at UAIC-IIT

A vision: our lives are time yards…

- We live our lives sequentially
  - *I lived within my family until the age of 16, then for two years I went in another city to follow a high school, then for 5 years I was a student in Bucharest, and one year before finishing the faculty I was married, then I came to Iași and 4 years later my first dauther was born, and when she was 5 I got divorced, one year later I married one more time...*

Our TYs

- Our lives meet, then split, then meet again…
- We meet many people during our lives, each having her/his own trajectory

=> like railways...

When I say:

I lived within my family until the age of 16…

- Actually, *each day I went to school, in the holidays I went to my grandparents house in the country side, and I went many times to cinema…*
- … and my parents were not near me all the time.

  => looked at it under the lens, a time segment reveals many more (finer) segments

  => it also depends on the **perspective**…

A Time Track perspective

- Somebody thinks/believes that somebody else has a certain trajectory

  => courses of events can be seen differently through different perspectives

  => sometimes TTs of a community of people are grossly seen intertwined, just like making a single destiny: collectivities, people living in a country

  => although each has her/his own (developing) TT

Just imagine…

- … that there is a device that records the TT of each one of us

  => technologically feasible

  $\Rightarrow$ although not very much wanted… (sentiment of BB oppression, leaks, etc.)

  $\Rightarrow$ but public lives of political personalities – on the internet

- THEN...

  => pages of history can be automatically written, with a proper intelligent summarisation tool

  => one can leave parts of her/his life to the posterity...

Acknowledgements

- Humanitas Publishing House for the Romanian version of Tash Aw's novel
- The Wylie Agency, UK, for the accept to use for research the original English version of Tash Aw's novel
- My classes of master students in CL, for debates while establishing the annotation conventions and participation in the first annotation experiment
- PhD student Andreea Macovei - for annotating the novel and co-authoring parts of this research
- Emi Peia for implementing an annotator and TY visualiser
- PhD student Șerban Boghiu – for continuing the research