Towards Compositional Semantics and Inference System for Telicity

Hitomi Yanaka
University of Tokyo
hyanaka@is.s.u-tokyo.ac.jp
LACompLing2021, December 16

Overview

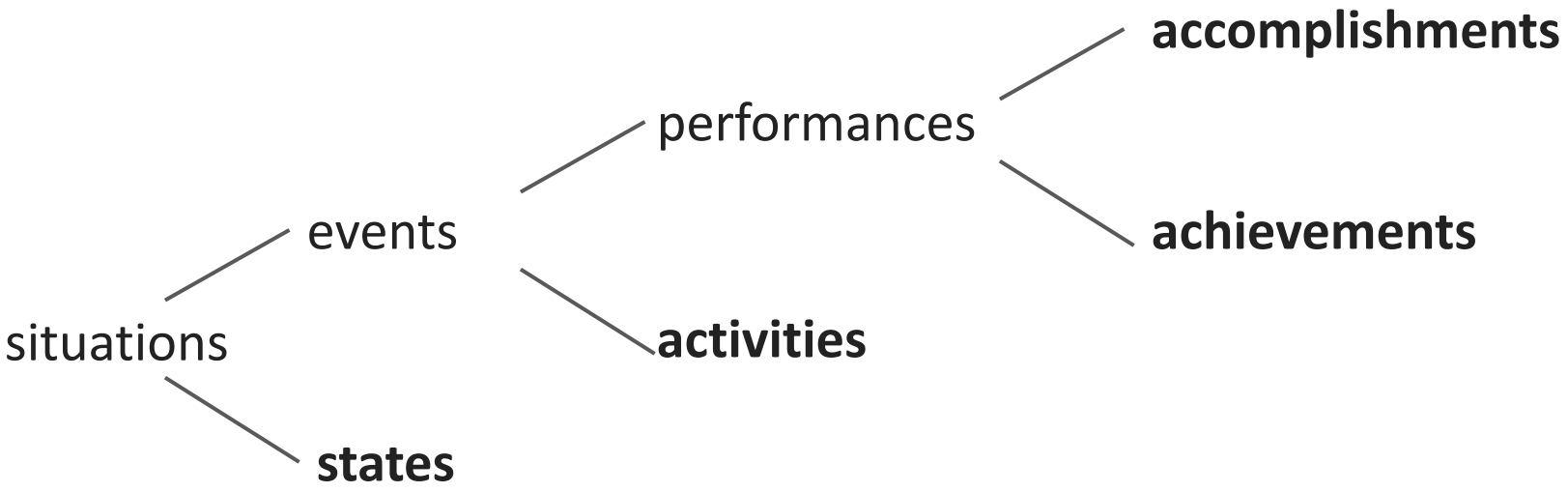
- Aspectual class of verbs: key role in tense and aspect
- Previous work on aspectual class of verbs:
  - Theoretical Linguistics[Vendler1957][Mourelatos1978][Dowty1979][Krifka1998]
  - Computational Linguistics[Friedrich and Palmer2014][Kober+2020]

Question: How to handle inference with aspectuality in computational linguistics?

➡ provide a hybrid approach of **distributional semantics** and **event semantics**

Aspectual class of verbs

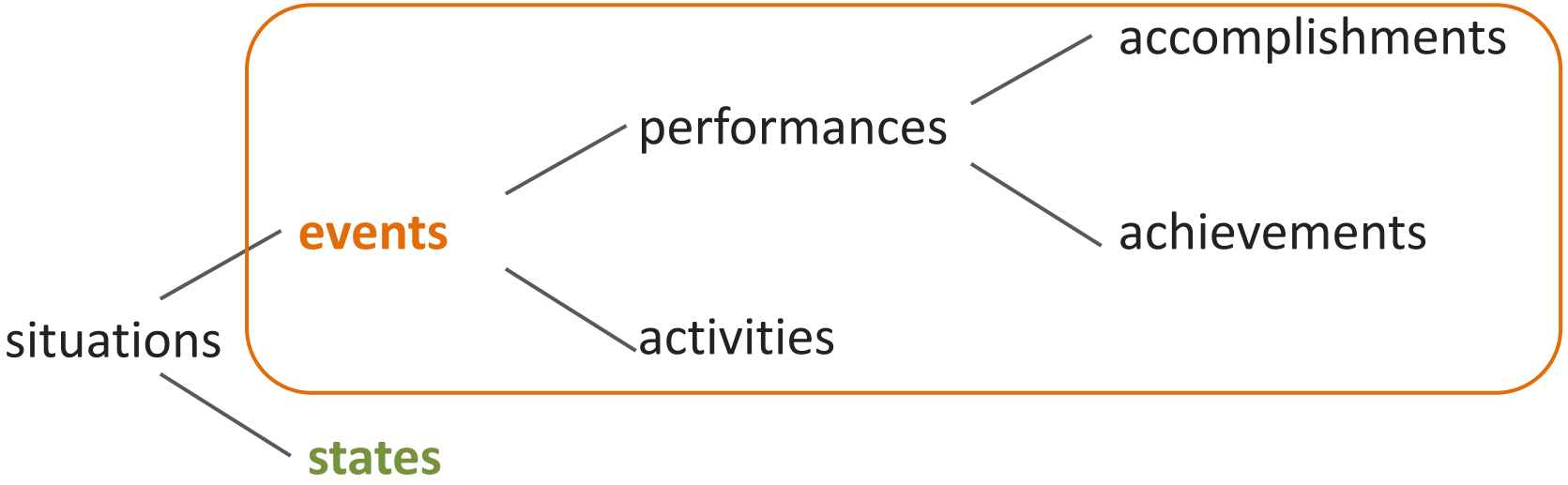
- semantic property of verbs that plays a key role in interpreting temporal structure[Moens and Steedman1988]
- typically categorized into 4 types[Vendler1957][Mourelatos1978][Bach1986]



Aspectual class of verbs: Event vs. States

- semantic property of verbs that plays a key role in interpreting temporal structure[Moens and Steedman1988]
- typically categorized into 4 types[Vendler1957][Mourelatos1978][Bach1986]
- main distinction 1:

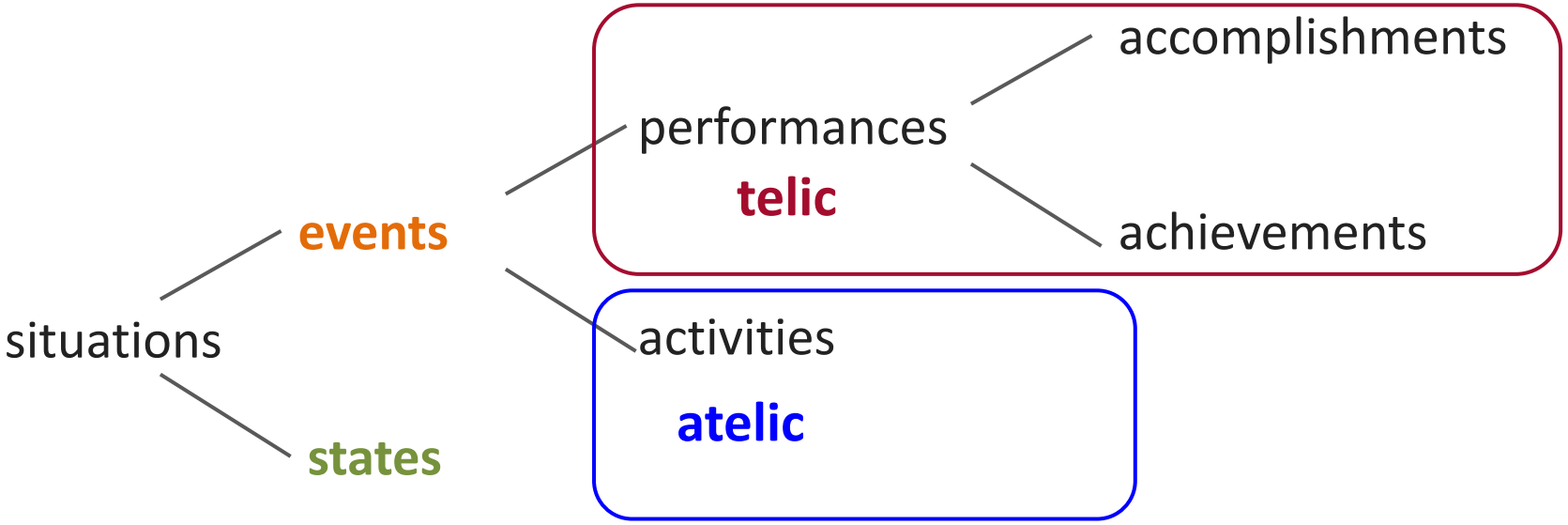
  events (e.g., run, draw a circle) vs. **states** (e.g., know, love)



Aspectual class of verbs: Telic vs. Atelic

- semantic property of verbs that plays a key role in interpreting temporal structure[Moens and Steedman1988]
- typically categorized into 4 types[Vendler1957][Mourelatos1978][Bach1986]
- main distinction 2:

  telic vs. **atelic** event (whether an event has a particular endpoint)



Imperfective Paradox[Dowty1979]

A telic event does not license entailment from its progressive form to the corresponding non-progressive form

Telic event:

T: John was writing a report ⇏ H: John wrote a report

Atelic event:

T: John was running in the park ⇒ H: John ran in the park

Imperfective Paradox[Dowty1979]

A telic event does not license entailment from its progressive form to the corresponding non-progressive form

Telic event:

T: John was writing a report ⇏ H: John wrote a report

Atelic event:

T: John was running in the park ⇒ H: John ran in the park

Question: How to handle inference with imperfective paradox in computational linguistics?

Previous work on compositional semantics in computational linguistics

Logic-based inference systems

- First-order prover and model builder[Bos and Markert 2005]
- Tableau prover for natural logic [Abzianidze2015 and 2016]
- Higher-order logic inference[Mineshima+2015][Martinez-Gomez+2017]
- Grammatical Framework and Coq-based inference system[Bernardy and Chatzikyriakidis2017 and 2021]

Computationally realize compositional semantics studied in formal semantics by defining lexical entries

Challenge: the telicity of verbs is complex

The telicity is not just lexically determined by a verb, but interacts with other words in a sentence[Bach1986][Krifka1998]

(1a) John wrote	**atelic**
(1b) John wrote a report	**telic**
(2a) John drank some coffee	**atelic**
(2b) John drank two cups of coffee	**telic**

How to computationally handle the telicity of verbs is a non-trivial issue

Previous work on aspectual class in computational linguistics

Various machine learning approaches for classifying aspectual class:

- Linguistic indicators for lexical aspects
  [Klavans and Chodorow1992][Siegel and McKeown2000]
- Cross-lingual projections[Friedrich and Gateva2017]
- Word representations using a distributional semantic model
  [Friedrich and Palmer2014][Friedrich+2016][Kober+2020]
  About 60-85% accuracy for the classification of clause-level aspects

Neural network-based Natural Language Inference

- BERT[Devlin+2019], RoBERTa[Liu+2019]
  Neural models do not perform aspectual inference very well
  [Kober+2019]

Hybrid approach of distributional semantics and event semantics

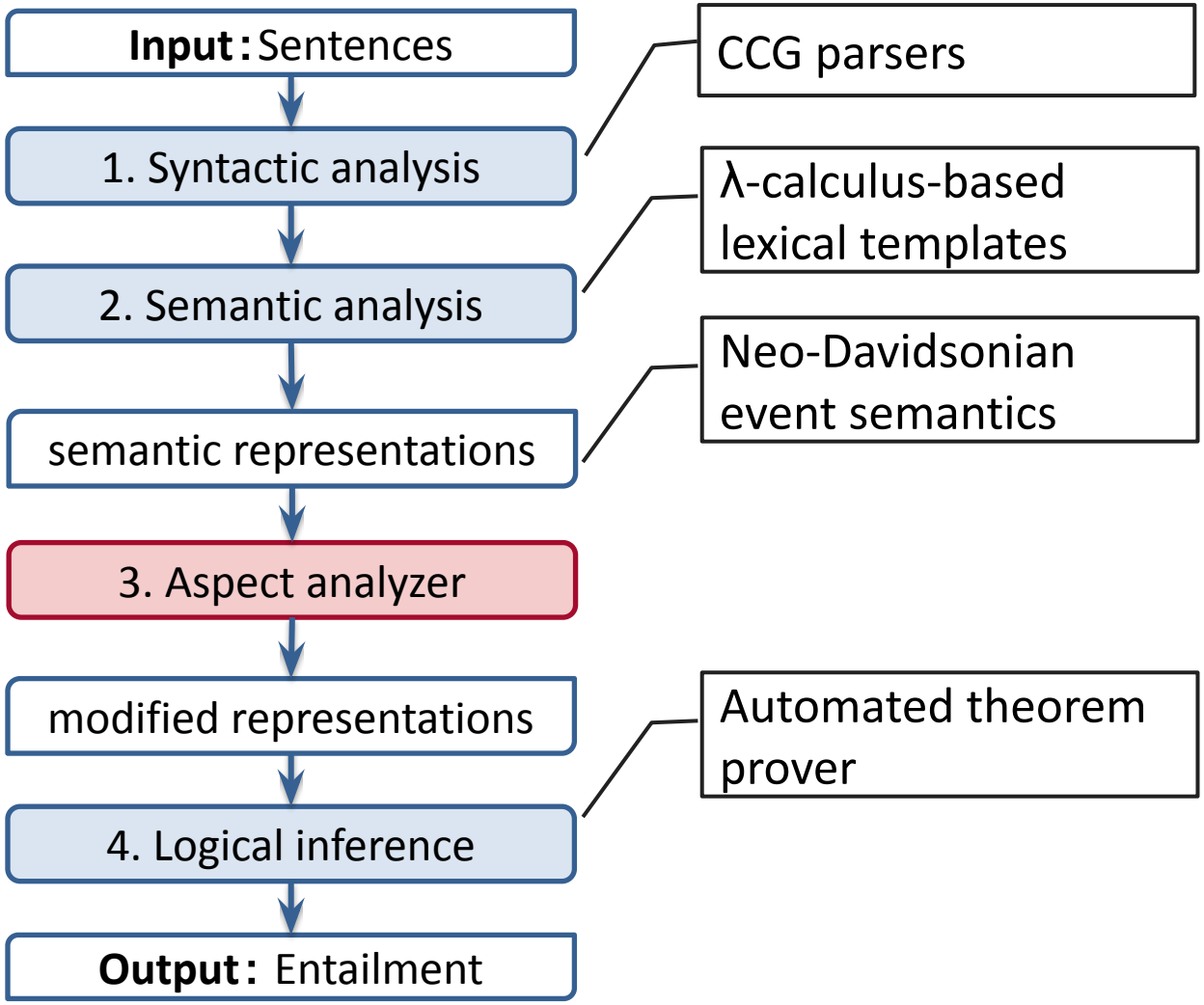
- Aim:
  1. a compositional semantics to map the aspectual class of verbs to semantic representations

  2. an inference system that performs inference with imperfective paradox

- Key idea:
  Combine **distributional semantics** with **event semantics** to handle inference with imperfective paradox

Overview of our proposed system

ccg2lambda
[Mineshima+,2015]
[Martinez-Gomez+,2016]



1. Syntactic analysis

Combinatory Categorial Grammar (CCG) [Steedman1996]

- Lexicalized grammar with clear syntax-semantics
- Robust CCG parsers trained on CCGBank[Hockenmaier and Steedman 2007]
  C&C[Clark and Curran2007], EasyCCG[Lewis and Steedman2014], depccg[Yoshikawa+2017]
- Lexical template:
  syntactic category and semantic representation for a word

Example: Some cats ran

$$\dfrac{\dfrac{\dfrac{\text{Some}}{NP/N}\quad\dfrac{\text{cats}}{N}}{NP}<\quad\dfrac{\text{ran}}{S\backslash NP}}{S}>$$

2. Semantic analysis

Neo-Davidsonian Event Semantics [Parsons1990]

- Every verb is decomposed into a predicate over events and a set of functional expressions relating the events
- **Event time** is a 1-place predicate over events **Dur(e)**
- **Temporal relation** between speech time (*st*) and reference time/ between event time and reference time[Reichenbach1956] Before/Equal/Meet/Overlap/Start/Finish/During [Allen1983]
- **Culmination point** is a 1-place predicate over events **Cul(e)**

Example: Some cats ran

$$\dfrac{\dfrac{\dfrac{\text{Some}}{NP[nb]/N} \quad \dfrac{\text{cats}}{N}}{\lambda F1F2.\exists x.(F1(x) \wedge F2(x)) \quad \lambda x.\text{cat}(x)}}{\lambda F2.\exists x.(\text{cat}(x) \wedge F2(x))}\,{<} \quad \dfrac{\dfrac{\text{ran}}{S[dcl]\backslash NP}}{\lambda Qt1.Q(\lambda x.\exists et2.(\text{run}(e) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(t1, t2) \wedge \text{Cul}(e) \wedge (\text{Subj}(e) = x)))}\,{>}$$

$$\dfrac{}{\exists x.(\text{cat}(x) \wedge \exists et2.(\text{run}(e) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2) \wedge \text{Cul}(e) \wedge (\text{Subj}(e) = x)))}$$

$$\exists x.(\text{cat}(x) \wedge \exists et2.(\text{run}(e) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2) \wedge \text{Cul}(e) \wedge (\text{Subj}(e) = x)))$$

3. Aspect analyzer

Our aspect analyzer is composed of three components:

3-1. Event extraction

Extract subformulas related to events from semantic representations

3-2. Event classification

Classify aspectual class of extracted events

3-3. Semantic recomposition

According to the predicted aspectual class, compose semantic representations again

Example demonstration (Telic event)

1. 2. Obtain tentative semantic representations via syntactic and semantic analysis

- In this step, we do not represent whether the event includes a culmination point

Telic event:

T1: John was writing a report ⇏ H1: John wrote a report

Semantic representation of T1:

$$\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{During}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2))))$$

Semantic representation of H1:

$$\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2))))$$

3-1. Event extraction

Extract a set of predicates related to events

- extract a set of predicates involving event variables
- extract a set of predicates which are accusative/dative cases of events

Telic event:

T1: John was writing a report ⇏ H1: John wrote a report

Semantic representation of H1:

$$\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2))))$$

Then, extract surface forms related to a set of predicates from an original sentence:

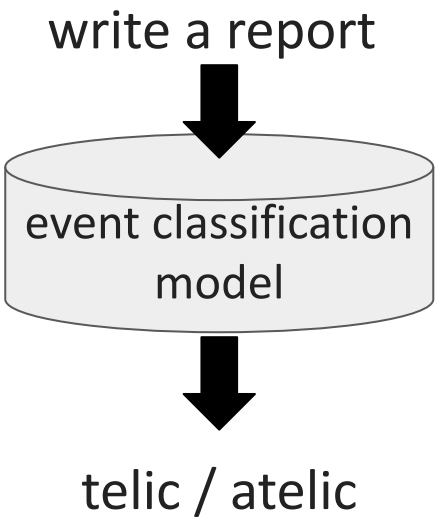
➡ write a report

3-2. Event classification

Classify the aspectual class (telic/atelic) of the extracted event phrase by a distributional semantic model and annotate **aspectual class tags**

- Previous work[Kobers+,2020] has reported that the accuracy with a verb phrase (local context) information is better than the accuracy with only a verb/sentence

Classify the aspectual class:



e.g., sequence classification using BERT[Devlin+2019]

Annotate aspectual class tags:

TEL
John wrote a report

ATE
John wrote

3-3. Semantic recomposition

According to the annotated aspectual class tag, compose semantic representations again

- In this step, we represent whether or not the event includes a culmination point
- The predicate Cul(e) is added only if the Part-of-speech tag is VBD (verb past tense) and the aspectual class tag is TEL (telic)

Lexical template examples

```
category: (S|NP)|NP
semantics: \E Q1 Q2 K t1. Q2(\x.True, \x.Q1(\y.True, \y.exists e t2.(E(e) &
Equal(t2,Dur(e)) & Before(t1,t2) & (Subj(e) = x) & (Acc(e) = y) & K(e,t2)),t1),t1)
pos: VBD
atag: TEL
```

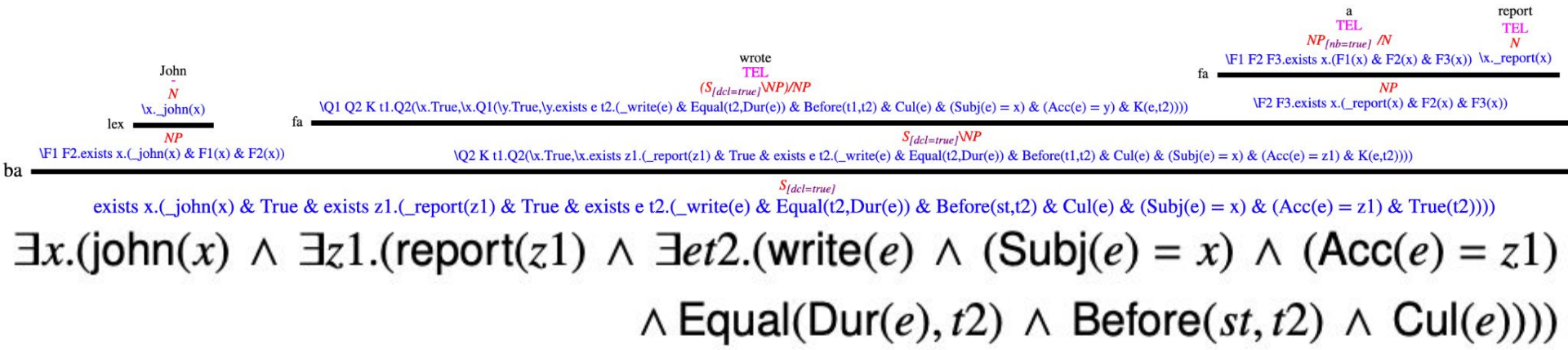
```
category: (S|NP)|NP
semantics: \E Q1 Q2 K t1. Q2(\x.True, \x.Q1(\y.True, \y.exists e t2.(E(e) &
Equal(t2,Dur(e)) & Before(t1,t2) & Cul(e) & (Subj(e) = x) & (Acc(e) = y) &
K(e,t2)),t1),t1)
pos: VBD
atag: ATE
```

3-3. Semantic recomposition: example

According to the annotated aspectual class tag, compose semantic representations again

- In this step, we represent whether or not the event includes a culmination point

Telic event: John wrote a report
TEL



$$\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2) \wedge \text{Cul}(e))))$$

4. Logical inference

- Convert semantic representations to typed First-Order forms[Sutcliffe2017]
- Try to prove the entailment relation by using the theorem prover Vampire
- Use axioms for First-Order Theory of Allen's Interval Algebra [Allen1983][Grüninger and Li,2017]

Telic event:

T: John <u>was writing a report</u> ↛ H: John <u>wrote a report</u>

TEL **TEL**

- We solve imperfective paradox by the existence of the predicate **Cul(e)**

T: $\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2))))$

↛ H: $\exists x.(\text{john}(x) \wedge \exists z1.(\text{report}(z1) \wedge \exists et2.(\text{write}(e) \wedge (\text{Subj}(e) = x) \wedge (\text{Acc}(e) = z1) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge \text{Before}(st, t2) \wedge \text{Cul}(e))))$

(*st*: speech time)

4. Logical inference

- Convert semantic representations to tptp format[Sutcliffe2017]
- Try to prove the entailment relation by using the theorem prover Vampire
- Use axioms for First-Order Theory of Allen's Interval Algebra
  [Allen1983][Grüninger and Li,2017]

Atelic event:

T: John was writing → H: John wrote

ATE **ATE**

- We solve imperfective paradox by the existence of the predicate **Cul(e)**

T: $\exists x.(\text{john}(x) \wedge \exists et2.(\text{write}(e) \wedge \text{During}(\text{Dur}(e), t2) \wedge (\text{Subj}(e) = x) \wedge \text{Before}(st, t2)))$

H: $\exists x.(\text{john}(x) \wedge \exists et2.(\text{write}(e) \wedge \text{Equal}(\text{Dur}(e), t2) \wedge (\text{Subj}(e) = x) \wedge \text{Before}(st, t2)))$

(st: speech time)

Discussion

Quantized property and cumulative property for representing the telicity of objects and events[Krifka1998][Zucchi&White2001][Rothstein2004]

- More fine-grained analysis for the telicity
- When we consider implementing Krifka's analysis computationally, distributional semantics should be also applicable to distinguish whether a predicate is quantized or cumulative

John drank a **quantity** of milk	**atelic**
John drank a **cup** of milk	**telic**
John drank a **cup** of milk **every day**	**habitual**

Conclusion and Future Work

- Aim:
  1. a compositional semantics to map the aspectual class of verbs to semantic representations
  2. an inference system that performs aspectual inference

- Key idea:
  - Combine **distributional semantics** with **event semantics**
  - Provide an aspect analyzer to represent the telicity of events

- Future work:
  - Cover various temporal and aspectual examples (e.g., temporal adverbials, habituality, states which are negative events)
  - Create a temporal and aspectual inference dataset and evaluate the proposed system

Thanks for listening!

Acknowledgements and References 1

Acknowledgements

We would like to thank two reviewers and Prof. Daisuke Bekki and Prof. Koji Mineshima for helpful comments. This work was partially supported by PRESTO, JST Grant Number JPMJPR21C8, Japan.

References

Zeno Vendler. 1957. Verbs and times. Linguistics in Philosophy

Alexander Mourelatos. 1978. Events, processes, and states. Linguistics & Philosophy

David R. Dowty. 1979. Word Meaning and Montague Grammar. Reidel

Annemarie Friedrich, Alexis Palmer. 2014. Automatic prediction of aspectual class of verbs in context. ACL

Thomas Kober, Malihe Alikhani, Matthew Stone, Mark Steedman. 2020. Aspectuality Across Genre: A Distributional Semantics Approach. COLING

Marc Moens, Mark Steedman. 1988. Temporal Ontology and Temporal Reference. Computational Linguistics

Judith L. Klavans, Martin Chodorow. 1992. Degrees of Stativity: The Lexical Representation of Verb Aspect. COLING

Eric V. Siegel, Kathleen R. McKeown. 2000. Learning Methods to Combine Linguistic Indicators:Improving Aspectual Classification and Revealing Linguistic Insights. Computational Linguistics

Annemarie Friedrich, Alexis Palmer, Manfred Pinkal. 2016. Situation entity types: automatic classification of clause-level aspect. ACL

References 2

Annemarie Friedrich, Damyana Gateva. 2017. Classification of telicity using cross-linguistic annotation projection. EMNLP
Thomas Kober, Sander Bijl de Vroe, and Mark Steedman. 2019. Temporal and aspectual entailment. IWCS
Johan Bos, Katja Markert. 2005. Recognising Textual Entailment with Logical Inference. EMNLP
Lasha Abzianidze. 2015. A Tableau Prover for Natural Logic and Language. EMNLP
Koji Mineshima, Pascual Martínez-Gómez, Yusuke Miyao, Daisuke Bekki. 2015. Higher-order logical inference with compositional semantics. EMNLP
Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL
Mark Steedman. 1996. Surface Structure and Interpretation. Linguistic Inquiry Monographs
Stephen Clark, James R. Curran. 2007. Wide-Coverage Efficient Statistical Parsing with CCG and Log-Linear Models. Computational Linguistics
Mike Lewis, Mark Steedman. 2014. A* CCG Parsing with a Supertag-factored Model. EMNLP
Terence Parsons. 1990. Events in the Semantics of English, A Study in Subatomic Semantics. MIT press
Manfred Krifka. 1998. The Origins of Telicity, Events and Grammar. Springer Netherlands