Strings for Temporal Annotation and Semantic Representation of Events

by

David Woods

A dissertation submitted

in fulfillment of the requirements

for the Degree of

Doctor of Philosophy

University of Dublin, Trinity College

November 2020

Declaration

I, the undersigned, declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I, the undersigned, agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

David Woods

November 2020

Abstract

Acknowledgements

This research is supported by Science Foundation Ireland (SFI) through the CNGL Programme (Grant 12/CE/I2267) in the ADAPT Centre (`https://www.adaptcentre.ie`) at Trinity College Dublin. The ADAPT Centre for Digital Content Technology is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

Related Publications

Contents

Declaration i

Abstract ii

Acknowledgements iii

Related Publications iv

1 Introduction 1

2 Relevant Literature 2

2.1 Times and Events . . . 2

2.1.1 Allen Relations . . . 2

2.1.2 Tense and Aspect . . . 2

2.2 Annotation . . . 2

2.2.1 ISO-TimeML . . . 2

2.2.2 TimeBank . . . 2

2.3 Semantics . . . 2

2.3.1 Discourse Representation Theory . . . 2

2.3.2 Boxer . . . 2

3 Finite State Temporality 3

3.1 Strings for Times and Events . . . 3

3.1.1 Creating Strings . . . 3

3.1.2 Granularity: Points vs Intervals vs Semi-intervals . . . 3

3.1.3 String Operations . . . 3

3.2 Applications . . . 3
3.2.1 Timelines from Texts . . . 3
3.2.2 Scheduling (Zebra Puzzle) . . . 3

4 Methods **4**
4.1 Extracting Strings from Annotated Text . . . 4
4.1.1 TLINKs . . . 4
4.2 Enhancing Strings using DRT . . . 4
4.2.1 Parallel Meaning Bank . . . 4
4.2.2 VerbNet and WordNet . . . 4
4.3 Inference via Residuals . . . 4

5 Implementation **5**
5.1 Computational Pipeline . . . 5

6 Evaluation **6**
6.1 Timeline Validity . . . 6
6.2 FRACAS Semantic Test Suite . . . 6
6.3 Correctness of Code . . . 6

7 Conclusion **7**

Bibliography **8**

Appendices **9**
Python Code . . . 9

1 Introduction

2 Relevant Literature

2.1 Times and Events

2.1.1 Allen Relations

2.1.2 Tense and Aspect

2.2 Annotation

2.2.1 ISO-TimeML

2.2.2 TimeBank

2.3 Semantics

2.3.1 Discourse Representation Theory

2.3.2 Boxer

3 Finite State Temporality

3.1 Strings for Times and Events

3.1.1 Creating Strings

3.1.2 Granularity: Points vs Intervals vs Semi-intervals

3.1.3 String Operations

3.2 Applications

3.2.1 Timelines from Texts

3.2.2 Scheduling (Zebra Puzzle)

4 Methods

4.1 Extracting Strings from Annotated Text

4.1.1 TLINKs

4.2 Enhancing Strings using DRT

4.2.1 Parallel Meaning Bank

4.2.2 VerbNet and WordNet

4.3 Inference via Residuals

5 Implementation

5.1 Computational Pipeline

6 Evaluation

6.1 Timeline Validity

6.2 FRACAS Semantic Test Suite

6.3 Correctness of Code

7 Conclusion

Bibliography

Appendices

Python Code