|  |
| --- |
| Recognising Temporal Expressions in Text |
| INTERIM REPORT  COM6920 Thesis Preparation |
| Christopher Northwood  *Supervisor: Mark Hepple* |

Table Of Contents

[Table Of Contents 3](#_Toc259525626)

[1 Introduction 4](#_Toc259525627)

[2 Background 5](#_Toc259525628)

[2.1 Temporal Expressions 5](#_Toc259525629)

[2.2 GUTime 5](#_Toc259525630)

[3 Work Plan 6](#_Toc259525631)

[3.1 Overview 6](#_Toc259525632)

[3.2 Python Tagger 6](#_Toc259525633)

[3.3 Automatic Rule Discovery 6](#_Toc259525634)

[3.4 Evaluation 6](#_Toc259525635)

[3.5 Timescale 7](#_Toc259525636)

[4 Bibliography 8](#_Toc259525637)

# Introduction

... maybe a page and a half

# Background

## Temporal Expressions

Background information about temporal expressions here (about 2/3rds page). The things themselves (events, and the canonical value) and relations between them. Talk about utility of this, why it’s desirable from a question answering pov, etc

Schilder, Katz, etc, contains some stuff about annotation. Ahn, Rantwijck blah, is probably better. Verhagen section 1 contains a good high level overview of temporal relations, with examples, etc

Relations sometimes necessarily for disambiguation/normalisation (Mazur/Dale). Recognition = discovery; normalisation = turning into ISO thingy (read up on standard to see how relative dates are discovered). Ahn, adafre, talks about separating recog/normalisation

### TimeML

More specific information about the TimeML standard here (another 2/3rds of page)

## Annotating Temporal Expressions

This is probably going to be the bulk of the report, a couple of pages maybe? Talk about hand tagging and introduction to automated tagging techniques. Talk about current state of the art stats (tarsqi paper comes up with some stuff for GUTime).

Setzer and gauzaukas talk about hand tagging, and the issues. There are many annotation guidelines. We’re using TimeML 1.2.1

### Rule-Based Tagging

Automated annotation based on hand-written rules

* Mani and Wilson 2000 (hand-written rules for expression tagging) – tempex (basis of GUtime)
* Tarqas (as a component)
* Tarsqi (uses GUTime as a component)
* DANTE
* Chronos
* Saquete
* Mikheev, et al; Krupka, Hausman(early work)

### Machine Learning

Automated annotation based on machine learning

Ahn, Rantwijck blah includes a particular architecture for machine learning which may be useful

* Mani and Wilson 2000 also includes a discussion of machine learning
* Kolomiyets...
* Hacioglu – SVM
* Baldwin
* Jang, Baldwin, Mani

...somehow fit information about other taggers here. TARSQI, has components, etc

### GUTime

Talk about GUTime specifically

## Evaluating Tagger Performance

Talk about TERN (need TERN dataset), TimeBank (have 1.1 – need 1.2), TempEval, AQUAINT. TempEval is for machine learning ones, I think?? (ask Mark for any relevant papers?)

# Work Plan

## Overview

High level work plan (introduction to different components, etc) about half a page

## Python Tagger

### Overview

Talk about how we’re going to port GUTime, etc

### Workload

|  |  |
| --- | --- |
| Task Name | Estimated Workload |
|  |  |
|  |  |
|  |  |
|  |  |

## Automatic Rule Discovery

### Overview

Talk again (about 2/3rds of a page) about how this is going to work

### Workload

|  |  |
| --- | --- |
| Task Name | Estimated Workload |
|  |  |
|  |  |

## Evaluation

### Overview

Talk about how the evaluation of my system is going to work (basically, as a TERN project)

### Workload

|  |  |
| --- | --- |
| Task Name | Estimated Workload |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |

## Timescale

# Bibliography

**There are no sources in the current document.**