A Scheme and Corpus for Classifying Temporal Orientation in Political Speech

Oscar Stuhler (oms279), Tirupal Rao Ravilla (trr321), Mark Jason Estudillo (mje341), Hee Seok Joo (hsj278)

Natural Language Understanding (DS-GA 1012), Spring 2019
Sam Bowman, Katharina Kann

Abstract

Studying temporal orientation has a long tradition in fields like psychologists or sociologists. While NLP research has undertaken considerable efforts regarding tasks like temporal scope annotation and relation extraction, as manifested in SemEval and TimeBank frameworks, relatively little work has been done on extracting the temporal orientation of speakers. In this paper, we (1) introduce a classification scheme for temporal orientation that comes with an elaborate annotation guide, (2) present a new corpus of 3000 sentences from the State of The Union Address (SOTU) annotated according to our scheme, (3) explore the utility of different model architectures for time orientation classification, and (4) investigate the temporal orientation in the SOTU from 1900 - 2019, showing that the current President Donald Trump has the lowest future orientation in modern American history.

1 Introduction

The study of temporal orientation has a long tradition throughout the social sciences. Psychologists have shown considerable interest in how human subjects experience time and how the latter affects their behavior (for a summary see Strathman & Joireman 2005). Among other things, temporal orientation has been shown to affect risk taking behavior (Boyd & Zimbardo 2005) abilities to cope with trauma (Holman & Cohen Silver 1998), health-related behavior (Crockett et al. 2009) While psychologists are mostly interested in temporal orientation as a cognitive property, work in other social science disciplines has investigated how pasts and futures are discursively constructed. A long standing tradition of research has investigated how stories about the past help to form what is referred to as collective memory and how collective memory is vital for the formation of group identities (Halbwachs 1950 [1992]; Olick et al. 2011). Narratives about futures, in turn, anchor decisions in markets (Beckert 2018), help to create political alliances (Mische 2013), motivate educational attainment (Frye 2012), and determine the fate of climate change politics (Hall 2016).

For all these different strands of research, the automatic extraction of temporal orientation from natural language is likely to be a major stepping stone. Against this backdrop, however, it is rather surprising that relatively little work in computational linguistics has tackled this task. While much effort has recently been devoted to annotating events with temporal scopes and extracting temporal relations between events (Verhagen et al. 2007, 2009, 2010; UzZaman 2013; Chang & Manning 2012; Vashishtha et al. 2019; Huang et al. 2016) only few studies have explicitly dealt with the task of extracting speakers' or writers' temporal orientation (Hasanuzzaman et al. 2017; Schwartz et al. 2015).

Our paper makes four contributions to temporal orientation extraction. Firstly, we introduce a classification scheme for temporal orientation that comes with an elaborate annotation guide, thereby advancing previous work that has relied on manually defining seed words (Hasanuzzaman et al. 2017) or ad-hoc coding (Schwartz et al. 2015). Secondly, we present a new corpus of 3000 sentences from the State of The Union Address (SOTU) annotated according to our scheme. Thirdly, we explore different approaches for time orientation classification: (1) We train a Support Vector Machine over an array of data representations: (a) a Bag-of-Words(BOW) model (b) a BOW with Parts of Speech (POS) tags and (c) a BOW with POS and bigrams (2) we use a stack of two bidirectional LSTMs over 300 dimensional GloVe embeddings, pretrained on Wikipedia and Gigaword. (3) we use a bi-directional gated RNN (GRU) with a logistic output unit (similar to Adel and Schütze (2017)) over 400 dimensional word2vec (Mikolov et al., 2013) embeddings, pretrained on Wikipedia, with and without attention to the hidden units. Our results suggest that classifying temporal orientation is tricky because of the lack of tense cues or common syntactic patterns in many sentences that could directly indicate a specific orientation. The RNN model outperforms all the other models. Finally, we use our data to estimate temporal orientation in the SOTU from 1900-2019. We demonstrate that there are considerable differences between American Presidents regarding their temporal orientation and that SOTU has become increasingly future oriented since 1900.

2 Previous Work in NLP

Studying temporal aspects of language is in no way new to NLP. Most of the existing research on time has focussed on the task of extracting the starting point of events and relating events temporally to each other. The Timebank corpus and TimeML scheme (Pustejovsky et al., 2003) provide an early general framework for temporal information extraction. More recently, three SemEval competitions (Verhagen et al. 2007, 2009, 2010; UzZaman 2013) have led to the creation of a large corpus that facilitates the evaluation of methods for temporal expression, event, and temporal relation extraction. The SemEval and the Timebank framework have in common that only a subset of event pairs are labeled with temporal relations. Cassidy et al. (2014) improved on this by introducing Timebank-Dense, a corpus where all edges between events in a sentence are labeled. Most recently, Vashishtha et al. (2019) expand on the work in the TempEval framework by introducing a new framework for extracting events' temporal relations and durations on the document level. A slightly different approach is taken by Huang and colleagues (2016), who introduced the EventStatus corpus, that cointrain a series of news coverage of civil unrest events. They tackle the task of distinguishing whether these events took place in the past, are still ongoing, or described as possibly taking place in the future.

The task that we and, we suspect, many social scientists, are interested in is in many regards much simpler than those tackled in the literature cited above: our goal is to extract the temporal orientation of a speaker or writer. While the literature above deals mostly with the extraction of precise time information for particular events, or temporal relations between different events mentioned in a sentence, we are interested measuring whether a speaker or writer references the past or on the future. To illustrate the difference of our task, consider the following two sentences:

- *Yes we can!*
- *Make America great again!*

While these two sentences provide no factual information about events in the world, the express two very different temporal orientations of the speaker. While the former is fundamentally about things happening in the future, the latter is more ambiguous. Similarly to the first sentence, the imperative "make" references things happening in the future, however, the word "again" unambiguously hinges at conditions of the past.

Work on temporal orientation in NLP has been surprisingly scarce. Schwartz et al. (2015) classify the temporal orientation of Facebook messages by classifying whether they reference the past, present, or future from the standpoint of the message sender. Hasanuzzaman et al. (2017) develop a temporal classification scheme for tweets that distinguishes between the categories past, present, future, and doubtful. While this is an important contribution, it leaves out the very real possibility that a tweet or sentence might reference the future as well as the past (in fact, 10.63% of the sentences in our corpus fall into this category). Hasanuzzaman and colleagues generate their training data by manually defining a vocabulary of seed terms that is expanded by adding the word2vec neighb of these terms. Training labels for their data are then generated by scoring tweets based on whether they contain these terms.

3 An Annotation Scheme for Temporal Orientation

Our classification scheme for temporal orientation is based on the sentence level. While some previous work on temporal orientation in NLP has relied on continuous scales of time (Schwartz et al. 2015), we argue that this makes for fuzzy boundaries and tough coding decisions. The distinctions we draw are simple, yet profound, and in the vast majority of the cases very intuitive to answer: firstly, does a sentence make a reference to the future? And secondly, does a sentence make a reference to the past? This leaves four possible labels for a sentence: {P:0; F:0}, {P:0; F:1}, {P:1; F:0}, and {P:1; F:1}. While clearly inferior in complexity to elaborate schemes like TimeML, we maintain that the simplicity of our scheme compared has the advantage of simple transferability. The costs associated with annotating even a small evaluation corpus in the TimeML scheme will make it not worthwhile to do so for most social scientists.

In some cases, past and future references are simply made by past or future tense use:

- *The program was improved at the last session of the Congress. {P:1; F:0}*
- *I am going to run for President. {P:0; F:1}*

But while the extraction of verb tenses is a fairly simple task, temporal orientation extraction is complicated by the fact that English offer many other ways to reference the past or the future.

- *We still have much to do. {P:0; F:1}*
- *History teaches us that strength is not simply a matter of arms. {P:1; F:0}*

A very common way to reference the future is the expression of hopes, goals, Often, this is achieved by the use of modal verbs, if clauses, or subjunctive mode. However, as the last example demonstrates, none of these features are a sufficient condition for future references:

- *We must strengthen the economy. {P:0; F:1}*
- *If water is cooled, it can turn to ice. {P:0; F:0}*

Importantly, we emphasize against previous work in NLP on temporal orientation that future and past references are independent of each other and can co-occur in a sentence:

- *We must now determine how we shall complete what we have begun.* {*P:1; F:1*}
- *We will always honor their memory.* {*P:1; F:1*}

An important aspect of our scheme is relativity, by which we mean that future and past are defined relative to the the perspective of the speaker and the moment of the speech act. Note that relativity implies, that there are instances in which future or past tense is used, but that neither contain a reference to the past nor to the future from the standpoint of the speaker:

- *He thought he was going to win, but his hopes were shattered.* {*P:1; F:0*}
- *The wife-beater, is inadequately punished by imprisonment: for imprisonment will often mean nothing to him, while it may cause hunger and want to the wife and children who have been the victims of his brutality.* {*P:0; F:0*}

A full explanation of our classification scheme and an annotation guide is provided in the supplementary materials.

4 Data

The State of the Union (SOTU) data utilized for analysis was obtained from The American Presidency Project (TAPP 2019). SOTU as an annual address made by the President of the United States to the American public. We decided to use speeches from US Presidents between 1900 and 2019. We omitted sentences before 1900 because of the stark changes in the syntactic structure of SOTU speeches from the earlier days. The data after 1900 has an average sentence length of 23.4 words.

We tokenized the speeches on the sentence level. The probability for a random sentence in our sample to contain a future reference was 56%. For past references, this probability is 39%. Lastly, data points that did not have a coherent sentences like 'IV' and 'IRRIGATION OF LAND' were not included as training data and were not annotated.

We post processed the data by tokenizing on the sentence level. We had three annotators, who annotated a total sample of 3000 sentences. 1500 of these sentences were cross annotated by two annotators. Inter annotator agreement is at .89 for past references and .87 for future sentences. Though this level of agreement is less than ideal, it is substantial enough to work with these labels. Because of time constraints, we did not get to resolve disagreements between annotators via deliberation. Instead, wherever there was disagreement between two annotators, we selected a random annotation.

5 Experimental Setup and Results

We evaluate on the data we annotated for two tasks: future orientation, and past orientation. Each task is a binary sentence classification task. For each sentence, the model has to decide whether it contains the respective temporal orientation. As the natures of the two tasks are very similar, we applied similar experimental analyses.

5.1 Baseline

Penn Treebank POS tags consist past and past participle tenses and modal verb tags. We used these tags to write some rules for each of the tasks. We run these rules to form a baseline model for our tasks. The results can be seen in Table 1 and Table 2.

5.2 Models

Firstly, We use a Support Vector Machine with a linear kernel (using scikit-learn, Pedregosa et al.) for both the tasks over different feature representation: BOW (one hot encoding on tokens), BOW with POS tags (one hot encoding on words and POS tags), BOW with POS tags and Bi-grams (one hot encoding on words, POS tags and word bi-grams).

Secondly, we apply a stack of two bi-directional LSTMs, with a fully connected hidden layer on top, as shown in Figure 1b. For input representations, we use 300 dimensional GloVe embeddings, pretrained on Wikipedia and Gigaword. In addition, we trained another LSTM model with the same architecture, but adding embeddings for the POS tags of the words (trained on the training dataset using word2vec) for input representations.

Lastly, we apply a bi-directional gated RNN (GRU - Cho et al., 2014) with an additional fully connected hidden layer, and logistic output unit (from Adel and Schütze(2017)). For the input representations, we use 400 dimensional word2vec embeddings, pretrained on Wikipedia, with an additional embedding for unknown words(⟨ unk ⟩). We further integrate an attention mechanism into the RNN model, on the hidden outputs of the RNN layer. Figure 1a gives an overview of how attention is applied to the RNN model being used.

Model	Accuracy(%)	P | R	F1
Baseline	63.41	0.77 | 0.63	0.58
SVM(BOW)	75.7	0.76 | 0.76	0.76
SVM(BOW+POS)	75.05	0.76 | 0.76	0.76
SVM(BOW+POS+Bigrams)	80.27	0.80 | 0.80	0.80
RNN	83.47	0.84 | 0.84	0.836
RNN with attention	**83.98**	0.83 | 0.85	**0.845**
LSTM	82.8	0.83 | 0.83	0.83
LSTM with POS	76.22	0.76 | 0.76	0.86

Table 1: Results for Future Orientation task

Model	Accuracy(%)	P | R	F1
Baseline	69.14	0.71 | 0.69	0.69
SVM(BOW)	73.01	0.74 | 0.73	0.72
SVM(BOW+POS)	75.54	0.76 | 0.76	0.75
SVM(BOW+POS+Bigrams)	78.4	0.79 | 0.78	0.78
RNN	79.08	0.88 | 0.63	0.73
RNN with attention	**80.44**	0.89 | 0.62	0.75
LSTM	78.75	0.79 | 0.79	**0.79**
LSTM with POS	72.34	0.78 | 0.72	0.70

Table 2: Results for Past Orientation task

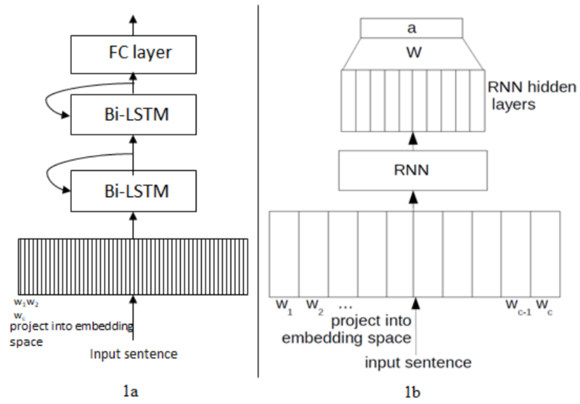

Figure 1.a. RNN network with attention on hidden representations; b. LSTM network

We primarily divided the data into training(80%), and testing(20%) datasets. Further, for hyperparameter tuning, we divided the training set into core training (75%) and validation set (25%.) . The results in the tables 1 and 2 give an overview of the performances of the techniques we explored. The baseline models perform better on the past task, while less so on the future task, because it is relatively easier to determine past tense from POS tags. All other models did better on the future task than on the past task. On the future task, RNN model with attention outperforms all the other models. On the past task, while RNN with attention does slightly better in terms of accuracy, LSTM and SVM do better in terms of F1.

Finally, we use our model to predict past and future references for all sentences in the post 1900 SOTU corpus. Our results suggest that references to the future have become much less likely since the 1980s. Current President Donald Trump continues this trend and has the lowest probability of making a future reference among all presidents. On the other hand, the probability of making a past reference seems to have no clear trend over time. Here, we see that Donald Trump has a much higher likelihood than the previous four Presidents.

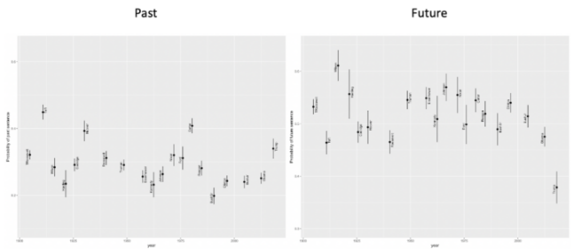

Figure 2. Past and Future Orientation Trend for Presidents

6 Discussion & Conclusion

Extracting temporal orientation from natural language is highly relevant to a variety of questions in the social sciences. In this paper, we presented an annotated corpus for classifying temporal orientation in political speech, an annotation scheme for the tasks, an analysis of how different NLP models perform on our corpus, and an investigation of the temporal orientation trends on post 1900 SOTU corpus. In light of our results from various models, it is evident that temporal orientation extraction is a non-trivial task. The complexity of many of the sentences in the dataset adds to the difficulty. And, although our dataset provides grounds to work on the problem, various factors like the variety of language style among different presidents, the multitude of ways of referring to past and future, a larger training corpus could improve the models' results for both tasks. That said, in light of the agreement levels between annotators, our models perform close to what could be reasonably expected. For future work, we hope to employ Active Learning techniques on the remainder of the SOTU addresses to utilize more data for training. Additionally, BERT (Devlin et al., 2018) has advanced the state-of-the-art in a number of NLP tasks; and since our task also relies on contextual relations between words, which BERT learns well, we hope to finetune and use BERT for our tasks. Furthermore, while we analyzed the SOTU addresses post 1900, the high average sentence length of the addresses before 1900 makes the task even more complicated for those earlier addresses. For future work, a major step forward would be to reduce the boundary to a more granular level and identify the scopes of temporal orientation markers.

Contribution Statement

All authors listed above contributed equally to this draft.

Code Repository

https://github.com/TirupalRavilla/temporal_orientation

References

Heike Adel and Hinrich Schütze. 2017. *Exploring different dimensions of attention for uncertainty detection*. In Proceedings of EACL.

Beckert, J. 2016. *Imagined Futures: Fictional Expectations and Capitalist Dynamics*. Cambridge, Harvard University Press.

Boyd, J. N. and P. G. Zimbardo. "Time perspective, health, and risk taking." In: Strathman, A. and J. Joireman, Eds. 2005. *Behavior in the Context of Time: Theory, Research, and Applications*. pages 85-107, Mahwah, Lawrence Erlbaum.

Cassidy, T., et al. 2014. An Annotation Framework for Dense Event Ordering *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Short Papers)*: 501-506.

Chang, A. X. and C. D. Manning. 2012. "SUTIME: A Library for Recognizing and Normalizing Time Expressions." *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC-2012)*: 3735-3740.

Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. 2014. Learning phrase representations using rnn encoder–decoder for statistical machine translation. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP).* pages 1724-1734, Doha, Qatar, October. Association for Computational Linguistics.

Crockett, R. A., et al. 2009. "Time orientation and health-related behaviour: measurement in general population samples." *Psychol Health* **24**(3): 333-350.

Frye, M. 2012. "Bright Futures in Malawi's New Dawn: Educational As- pirations as Assertions of Identity." *American Journal of Sociology* **117**(6): 1565-1624.

Halbwachs, M. 1992 [1950]. *On Collective Memory.* Chicago, University of Chicago Press.

Hall, J. R. 2016. "Social futures of global climate change: A structural phenomenology." *American Journal of Cultural Sociology* **4**(1): 1-45.

Hasanuzzaman, M., et al. 2017. Temporal Orientation of Tweets for Predicting Income of Users. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers):* 659-665.

Holman, A. and R. Cohen Silver. 1998. "Getting "Stuck" in the Past: Temporal Orientation and Coping With Trauma." *Journal of Personality and Social Psychology* **74**(5): 1146-1163.

Huang, R., et al. 2016. "Distinguishing Past, Ongoing, and Future Events: The EventStatus Corpus." *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing:* 44-54.

Olick, J. K., et al., Eds. 2011. Introduction. in: Olick et al (eds.): *The Collective Memory Reader.* Oxford, Oxford University Press.

Mische, A. 2014. "Measuring futures in action: projective grammars in the Rio+20 debates." *Theory and Society* **43**: 437-464.

Pustejovsky, J., et al. 2003. "The TIMEBANK Corpus." *Corpus linguistics:* 647-656.

Schwartz, A. H., et al. 2015. "Extracting Human Temporal Orientation from Facebook Language." *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language.*

Strathman, A. and J. Joireman, Eds. 2005. *Behavior in the Context of Time: Theory, Research, and Applications.* Mahwah, Lawrence Erlbaum.

UzZaman, N., et al. 2013. "SemEval-2013 Task 1: TEMPEVAL-3: Evaluating Time Expressions, Events, and Temporal Relations." *Second Joint Conference on Lexical and Computational Semantics:* 1-9.

Vashishtha, S., et al. 2019. "Fine-Grained Temporal Relation Extraction." *arXiv:1902.01390v1.*

Verhagen, M., et al. 2007. "SemEval-2007 Task 15: TempEval Temporal Relation Identification." *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007):* 75-80.

Verhagen, M., et al. 2009. "The TempEval challenge: identifying temporal relations in text." *Language Resources and Evaluation* **43**(2): 161-179.

Verhagen, M., et al. 2010. "SemEval-2010 Task 13: TempEval-2." *Proceedings of the 5th International Workshop on Semantic Evaluation, ACL 2010:* 57-62.

UzZaman, N., et al. 2013. "SemEval-2013 Task 1: TEMPEVAL-3: Evaluating Time Expressions, Events, and Temporal Relations." *Second Joint Conference on Lexical and Computational Semantics:* 1-9.

Annotation Guidelines

Our annotation scheme works on the sentence level. We have two binary classification categories that are independent of each other.

Past

Category one is question of whether the sentence contains a reference to the past. There are many instances in which the past is directly referenced by using the past tense.

- *The program was improved at the last session of the Congress.* {*P:1; F:0*}

However, there are many ways to refer to the past that do not make use of past tense such es in the example below:

- *The American history is a history of class struggle.* {*P:1; F:0*}

Future

Category two is question of whether the sentence is projective. A sentence is projective if it contains contains a reference to the future. In many cases, this is done by using future tense:

- *I am going to run for President.* {*P:0; F:1*}

Similarly to past references, however, many future references are made without future tense use. One way of doing so, is to use lexical features that mark projectivity:

- *Our objective in the world is peace.* {*P:0; F:1*}
- *We are prepared to devote our energy and our resources to this task.* {*P:0; F:1*}
- *We know that our own security and the future of mankind are at stake.* {*P:0; F:1*}
- *We are working hard at that task.* {*P:0; F:1*}
- *This is where America needs to go.* {*P:0; F:1*}
- *We still have much to do.* {*P:0; F:1*}
- *There are some clouds on the horizon.* {*P:0; F:1*}
- *One of America's most important institutions is in need of reform.* {*P:0; F:1*}

While the first sentence can be read as a general claim, the keyword "objective" highlights its projectivity of this phrase. The second example similarly references the future without explicitly making use of future tense by using the phrase "prepared to devote". In the third example, the word "future" itself highlights its projectivity. In the 4th example, the speaker uses the present tense to refer to work that is being done, but he also makes explicit that there is an ongoing "task" that has yet to be completed, thereby referencing the future.

Another common way to reference the future is the use of modal verbs and if clauses or subjunctive mode:

- *We should really go home.* {*P:0; F:1*}
- *We shall continue to give our support to the United Nations.* {*P:0; F:1*}
- *We should strengthen the economy.* {*P:0; F:1*}
- *If we win the election, we might implement the policy.* {*P:0; F:1*}
- *If we won the election, we could implement the policy.* {*P:0; F:1*}

In the first example, the speaker uses "should," which is clearly aspirational and therefore projective. The same is true for the sentences that follow. The last two sentences provide two examples with projective if-clauses. Though note that not every use of an if clause is projective:

- *If water is cooled to 0 degree Celsius, it will turn to ice.* {*P:0; F:0*}

Just as much as the use of a modal verb does not necessarily make a future reference:

- *In two months, June and July, more than a thousand airplanes that could have been made and should have been made were not made.* {*P:1; F:0*}
- *Direct disbursements to the veteran have resulted, which otherwise would not have been made.* {*P:1; F:0*}
- *Many people watching tonight can remember a time when finding a good job meant showing up at a nearby factory.* {*P:1; F:0*}
- *This journey began 43 years ago, when a woman named Rosa Parks sat down on a bus in Alabama and wouldn't get up.* {*P:1; F:0*}

While modal verbs like 'should' and 'shall' are the most common ways to express the need for action and thereby make a future reference, this can also be achieved by other linguistic features:

- *I urge you to pass association health plans.* {*P:0; F:1*}
- *I am asking you for your support of this bill.* {*P:0; F:1*}
- *Our immediate action is required.* {*P:0; F:1*}
- *Let's strengthen the economy.* {*P:0; F:1*}

Relativity

That said, note that we mean future and past as as relative to the moment and perspective of the speaker and the speech act. The following example contains the two keywords "imagine" and "future:"

- *Before his hopes were shattered, he had imagined a great future for himself.* {*P:1; F:0*}
- *They planned for their future.* {*P:1; F:0*}

Note that relativity implies, that there could also be instances in which the future tense is used, but that do not contain a reference to the future from the standpoint of the speaker:

- *He thought he was going to win, but his hopes were shattered.* {*P:1; F:0*}

Finally, note that there are some cases where the use of past tense is not a reference to the past from the standpoint of the speaker. Similarly, there are cases where the use of future tense is not a reference to the future from the standpoint of the speaker:

- *The wife-beater, is inadequately punished by imprisonment: for imprisonment will often mean nothing to him, while it may cause hunger and want to the wife and children who have been the victims of his brutality.* {*P:0; F:0*}

Independence of Past and Future Classification

Note that these two categories are independent of each other. It should be obvious, that a sentence need not contain a reference to either the future or the past, but can purely express something about the current state of affairs:

- *We have a more productive economic system and a greater industrial potential than any other nation on the globe.* {*P:0; F:0*}
- *South Korean homes have greater Internet access than we do.* {*P:0; F:0*}

A special case are sentences that comprise general propositions. We treat such cases as neither containing a future or past reference unless they make the past or future reference explicit:

- *Strength is not simply a matter of arms.* {*P:0; F:0*}
- *These events teach us that strength is not simply a matter of arms.* {*P:1; F:0*}
- *It is our duty to defend our nation.* {*P:0; F:1*}

In the second sentence, the keyword duty highlights that this general proposition is immediately relevant to the speakers projected course of action.

Furthermore, note that there are sentences that simultaneously contain references to the past and to the future:

- *The first half of the century will be known as the most turbulent in history.* {*P:1; F:1*}
- *We must now determine how we shall complete what we have begun.* {*P:1; F:1*}
- *We will always honor their memory.* {*P:1; F:1*}
- *We have made a beginning in preservation, but it is only a beginning.* {*P:1; F:1*}
- *We have much at stake in avoiding a repetition of that calamity.* {*P:1; F:1*}
- *If anyone tells you America's best days are behind her, they're looking the wrong way.* {*P:1; F:1*}
- *We are approaching out third century of independence.* {*P:1; F:1*}
- *The process has begun.* {*P:1; F:1*}
- *We must not go back to the days of "the hollow army".* {*P:1; F:1*}
- *The cattle industry has not yet recovered.* {*P:1; F:1*}

In the last sentence, the word "yet" hinges at the future even though the sentence mainly describes the past.

General Statements and Ongoing Events

Two types of sentences are of worthy of special consideration in the annotation. One of them are sentences that make general truth claims, such as the following:

- *Productivity is the foundation-stone of our security structure.* {*P:0; F:0*}

Such a truth claim, as long as it is not bounded by a temporal qualification, by virtue of being general can be read as a claim about the past as well as the future. However, the sentence itself entails no explicit reference to the future or past. We take a conservative approach and code such statements as {P:0; F:0}, unless they make their future of past reference explicit by adding some cue. We do this for pragmatic reasons because otherwise one would have to define the boundary of what constitutes a general truth claim which in itself is very difficult and hard to implement in practice. We do, however, acknowledge that other conceptions of projectivity may warrant a different coding scheme. Furthermore, we acknowledge that the question of what makes a legitimate cue is subject to the annotators interpretation, and that, in some rare cases, these interpretations can differ among annotators. Below are some examples of sentences with general claims and how they were coded.

- *I believe good will must be dedicated to great goals.* {*P:0; F:1*}

Here, the speaker situates himself in the present to this truth claim. From this vantage point, the keyword "goals" appears to constitute a future reference.

- *The problem is that monopolistic activity prevents investment in new production facilities.* {*P:0; F:1*}

Similarly, in the example above, the use of "the problem" makes extremely likely, that the speaker takes about an acute problem, and therefore the phrase that follows constitutes a reference to the future. To make this point clear, consider the following sentence where "The problem" is replaced with "Our problem":

- *Our problem is that monopolistic activity prevents investment in new production facilities.* {*P:0; F:1*}

In this example the future reference is unambiguous because the speaker situates himself in the situation which implies that the sentence that follows must refer to the future. We interpret "The problem" in the first sentence to mean "Our problem" or "My problem," and therefore code it as future reference.

- *The use of detectives, though often necessary, tends towards abuse, and should be carefully guarded.* {*P:0; F:1*}

The second type of sentences that should be concerns sentences that describe ongoing events or condition. Below are a number of examples:

- *People are out of work.* {*P:0; F:0*}
- *Our plant productivity is increasing.* {*P:0; F:0*}
- *These conditions prevail today.* {*P:0; F:0*}

In all these instances, it appears that the speaker is describing a current event or condition and the nature of these events or conditions make plausible the inference that these are ongoing, that is, that they have been going on for at least some time in the past and will go on for at least some time in the future. However, there is nothing in the sentence that makes this fact explicit. Our approach is therefore similar to the one we follow for general truth claims. We code them as neither constituting a future or past reference, unless there are some additional elements of the sentence that makes this explicit, such as in the examples below:

- *Our plant productivity is not increasing fast enough.* {*P:0; F:1*}

By saying, that the "productivity" is not increasing fast enough, the speaker hinges at the future. He points out a need for change that is made necessary by projected events in the future.

- *Our efforts to reach agreements on peace settlements for these countries have so far been blocked.* {*P:0; F:1*}

Here, the speaker makes clear that "efforts" have been taken, which constitutes a past reference. However, the use of "so far" also hinges at the future, making it explicit that this might change in the future.

- *Work in canal construction has been in progress for less than a year.* {*P:1; F:0*}

Here, the speaker makes it clear that an event has been going on for some time and we might be led to infer that it is still ongoing. However, there is nothing in the sentence that makes this explicit.