A Multi-Pass Sieve Coreference Resolution for Indonesian

Valentina Kania Prameswara Artari, Rahmad Mahendra,
Meganingrum Arista Jiwanggi, **Adityo Anggraito**, **Indra Budi**
Faculty of Computer Science, Universitas Indonesia
Depok 16424, West Java, Indonesia
`valentina.kania@ui.ac.id, rahmad.mahendra@cs.ui.ac.id`

Abstract

Coreference resolution is an NLP task to find out whether the set of referring expressions belong to the same concept in discourse. A multi-pass sieve is a deterministic coreference model that implements several layers of sieves, where each sieve takes a pair of correlated mentions from a collection of non-coherent mentions. The multi-pass sieve is based on the principle of high precision, followed by increased recall in each sieve. In this work, we examines the portability of multi-pass sieve coreference resolution model to Indonesian language. We conduct the experiment on 201 Wikipedia documents and multi-pass sieve system yields 72.74% of MUC F-measure and 52.18% of BCUBED F-measure.

1 Background

Many Natural Language Processing (NLP) tasks need to incorporate linguistic comprehension beyond semantics understanding. Coreference resolution is an important discourse-level NLP pipeline that can be used to support a number of NLP applications, such as question answering, summarization, and dialogue system. Coreference resolution task aims to evaluate whether a set of expressions in the text refer to each other, in other words whether they describe the same entity in real-world situation (Hirschman and Chinchor, 1998; Sukthanker et al., 2020)

There are not many coreference resolution studies in Indonesian. Budi et al. (2006) worked on the Indonesian coreference resolution task by applying the Association Rules. Suherik and Purwarianti (2017) developed a coreference resolution system using supervised classifier. They utilized lexical and syntactic features to connects pronouns to named entities, between named entities, and between pronouns.

Stanford NLP lab introduced Multi-Pass Sieve approach for coreference resolution task (Raghunathan et al., 2010; Lee et al., 2011). Then, this method has been widely adapted for the same task in various languages other than English, such as Korean(Park et al., 2014), Basque (Soraluze et al., 2019), Arabic, and Chinese (Chen and Ng, 2012).

The Multi-Pass Sieve concept is based on the principle of high precision, followed by increased recall in each pass. In this approach, the breakdown of the correlation relationship is divided into several layers of sieves. In spite of emphasis on hand-crafted linguistics features, this is a robust technique to tackle the coreference resolution task. In addition, the results of this approach have a high level of interpretability, so that the analysis of the results has the potential to be used in long-term research.

In this research, we utilize this Multi-Pass Sieve technique for Indonesian coreference resolution task. In addition, we build a data set for testing the system. We publish our code and data for research purpose[1].

2 Noun Phrases in Indonesian

The set of referential phenomena in Indonesian includes several types, i.e., pronouns, demonstratives, noun phrases, and named-entities. In this section, we briefly describe the first three types of nominal, while there is no difference of names between Indonesian and English.

Pronouns Pronouns are a class of words that are used to refer to another noun (Alwi et al., 2010). A personal pronoun is a pronoun which refers to one or more persons. Personal pronouns in Indonesian are differentiated into singular and plural pronouns. Unlike English, pronouns in Indonesian are not

---

[1]`https://github.com/valentinakania/`
`indocoref`

differentiated by gender, nor by its function in a sentence, i.e., same set of pronouns are used as a subject, an object, or to indicate possession. For example, the word "*saya*" is a personal pronoun in Indonesian which can be used as a subject pronoun, e.g., "**Saya** *makan*" (in English "**I** *am eating.*"), an object pronoun, e.g., "*Pak Bob memanggil* **saya**." (in English "*Mr.Bob is calling* **me**.", and a possesive pronoun, e.g., "*Ini buku* **saya**." (in English "*This is* **my** *book.*").

Indonesian pronouns can be in form of the clitics (Larasati, 2012). A clitic is a morpheme that is attached to another word or phrase. Clitic pronouns in Indonesian include "-*ku*" as first-person pronouns, "-*mu*" as second-person pronouns, and "-*nya*" as a third-person pronoun.

Demonstratives Demonstrative, also classified as a demonstrative pronoun by Alwi et al. (2010), are words that refer to a noun or noun phrase. In Indonesian, there are two common words that are classified as demonstratives, namely "*ini*" (in English:"*this*" or "*these*") and "*itu*" (in English: "*that*" or "*those*").

Nouns Nouns can be seen semantically as words that represent humans, animals, objects, meanings, and concepts that exist in the world as described in Alwi et al. (2010). Syntactically, a noun can be negated with "*bukan*" (in English: "*not to be*"), but cannot be negated with "*tidak*" (in English: "*do not*"). Nouns can be followed by one or more adjectives, either directly or connected by the word "*yang*" (in English: "*which*"). Several Indonesian nouns are multiword expression (Suhardijanto et al., 2020), e.g., "*kamar tidur*" (in English: "*bedroom*") and "*rumah sakit*" (in English: "*hospital*")

Noun Phrases A noun phrase may consist of one or more noun(s), pronoun(s), numeral(s), verb(s), adjective(s), and demonstrative(s). A noun phrase in Indonesian is constructed by expanding the noun to the left, with a determiner, or to the right, with modifiers. The initial noun before the expansion is named as the core noun (the head word). While English put modifiers before the head word, modifiers come after the head word in Indonesian, e.g., "*buku matematika*" vs. "*mathematics book*".

Here are some rules for expanding a head word into a noun phrase in Indonesian (Alwi et al., 2010).

1. A head word can be expanded to the left with numerals or numeral phrases.

satu	meja
NUM	HEAD

"one table"

2. A head word can be expanded to the right by one or more other nouns (explanatory nouns/ EXPN), then followed by a personal pronoun (PP), then it can be followed by a demonstrative word (DEM).

meja	kayu	mereka	ini
HEAD	EXPN	PP	DEM

"this dining table of theirs"

3. A head word can be expanded to the right by zero or more adjectives, pronouns or pronominal phrases, followed by a demonstrative.

meja	biru	Ibu
HEAD	ADJ	PP

"mother's blue table"

3 Multi-Pass Sieve Coreference Model for Indonesian

We design multi-pass sieve to resolve coreference problem in Indonesian text. The Multi-Pass Sieve approach works by receiving input in the form of mention pairs $(m1, m2)$ and classifying the correlation relationship between the two mentions according to the definition of each sieve layer sequentially and stops when the pair $(m1, m2)$ has been declared to have a correlation relationship or when the pair have visited the last sieve. The model is implemented in six tiers, i.e., exact string match, precise constructs, string head match, proper head word match, relaxed head match, and pronoun sieve.

Several other sieves in original Stanford model (Lee et al., 2011) are not adapted in our model due to the differences of linguistic characteristics between English and Indonesian languages. In contrast to English, demonym relation in Indonesian are not expressed by changing words. For example: "*Indonesia*" and "*orang Indonesia*" (in English: "Indonesia" and "Indonesian"). A sieve-pass using morphological properties is not included in ours. Indonesian word does not possess gender, number (singular vs. plural), and animacy attribute.

On the other hand, lexical chain sieve is not implemented due to the lack of language resources. Existing Indonesian WordNet (Putra et al., 2008)

does have neither adequate synsets nor lexical semantic relation. In addition, we do not find proper dictionary to construct alias sieve. We also exclude relaxed string match as adverb clauses are out of scope of mention detected in our coreference resolution system. Discourse processing sieve is skipped since the data used in our experiment does not contain any direct sentence.

3.1 Pass 1: Exact String Match

In this first layer of our Multi-Pass Sieve model, each pair of noun phrases is considered coherent if the two strings are the same, regardless of the difference in capitalization and punctuation.

3.2 Pass 2: Precise Constructs

The second layer of the model determines two noun phrases as coreferent if they follow a certain language structure.

Appositive Relationship . It is indicated by the position of the two mentions in the sentence in which a mention appears next to another mention. The second mention serves to explain the first mention, which is a proper noun in most cases. Both phrases are usually separated by a punctuation, i.e., comma, semicolon, dash, or brackets.

> [Emma Pillsburry], [salah seorang guru di SMA Ohio], mendengar berita itu.
>
> "[Emma Pillsbury], [one of the teachers at Ohio High School], heard the news."

Copulative Relationship . The copulative relationship between two mentions is explanatory, in which both mentions are connected by a copula verb, e.g., *"adalah"*, *"merupakan"*, *"yakni"*, *"yaitu"*, and *"ialah"*, (in English: *"is/am/are/was/were"*). Like appositive relationship, a mention explains another mention in copulative relationship.

> [Feng Yuxiang] adalah [seorang panglima perang pada masa Republik Tiongkok awal abad-20].
>
> "[Feng Yuxiang] was [a warlord during the early 20th century Republic of China]."

Abbreviation Relationship . The relationship exists when a mention is an abbreviation of another mention. As shown in Budi et al. (2006) and Lee et al. (2011), the abbreviation is detected by using a set of patterns.

> [Pekan Olahraga Nasional] (disingkat [PON]) adalah pesta olahraga nasional di Indonesia yang diadakan setiap empat tahun sekali dan diikuti seluruh provinsi di Indonesia.
>
> "[Pekan Olahraga Nasional] (abbreviated [PON]) is a national sporting event in Indonesia which is held every four years and is attended by all provinces in Indonesia."

In aforementioned example, "PON" is detected as the abbreviation for noun phrase "Pekan Olahraga Nasional", by matching the first letter of each word in the phrase.

3.3 Pass 3: Strict Head Match

In the third layer of our Multi-Pass Sieve model, two noun phrases are evaluated as coreferent if the head word of both phrases are the same and they also share the same lexical class. There are two variations of the strict head match. (i) **Strict head match**, looks at the similarity of core nouns consisting of one word taken during data preprocessing. (ii) **Full head match**, sees the similarity of the core nouns consisting of several words. The choice of full head in this case is to input words with the POS Tag NOUN or PROPN as the head.

There are three passes that take advantage of the head word or core noun features, i.e., the demonstrative relationship, name abbreviation, and strict head match itself.

Demonstrative Relationship The demonstrative relationship between two noun phrases, apart from depending on the position of the phrase in the sentence, also uses demonstrative word classes to determine the relationship. Noun phrases A and B are said to have a demonstrative relationship if one of the phrases contains a demonstrative, and after removing the demonstrative word, the phrase is a sub-phrase of another phrase.

> "The Break Up" adalah [episode keempat dari serial televisi komedi musikal Glee musim keempat]. [Episode ini] diskenarioi oleh Ryan Murphy dan disutradarai oleh Alfonso Gomez-Rejon.
>
> "The Break Up" is [the fourth episode of the fourth season of the musical comedy television series Glee]. [This episode] was screenplayed by Ryan Murphy and directed by Alfonso Gomez-Rejon."

Short Name The short name feature sees if one mention is the short name of another mention. The implementation of this feature is as follows: given two noun phrases A and B, where B is shorter than A, B is the short name of A if every word in B is in A, and one of A or B is a proper noun. Examples of cases of short names are found in the mention of people using nicknames or last names in the article text, after the full name is mentioned at the beginning.

3.4 Pass 4: Proper Head Word Match

At this layer, the system looks for a specific core noun similarity for PROPN POS Tag. If given two noun phrases A and B which are identified as proper names, the system assesses the two corresponding noun phrases if A and B have core nouns with the same PROPN POS Tag and A and B have compatible attributes. The attribute used in this research is named-entity class.

> Will kemudian memberitahu tunangan-nya, [guru bimbingan konseling Emma Pillsbury], bahwa ia telah diterima di dewan seni pemerintah. [Emma] tidak mau meninggalkan Lima dalam waktu yang lama.
>
> "Will then told his fiancé, [counseling teacher Emma Pillsbury] that he had been accepted on the government arts council. [Emma] didn't want to leave Lima for a long time."

3.5 Pass 5: Relaxed Head Match

In this layer, given two noun phrases A and B, Relaxed Head Match compares whether each word in the head noun A is in the noun phrase B. The head noun A may consist of several words.

> Peternakan Nenek Bebek menjadi [pusat pertemuan keluarga], di mana [pertemuan tersebut] diatur oleh Nenek Bebek.
>
> "Grandma Duck's ranch became [the center of the family gathering], where [the meeting] was arranged by Grandma Duck."

3.6 Pass 6: Pronouns Sieve

The Pronoun layer works anaphorically. If the mention is in the form of pronouns, both words and clitics, the candidate antecedents may only be the noun phrases previously mentioned. In this study,

a candidate pronoun will be paired with the closest noun phrase that does not violate the following rules:

- Noun phrases identified as location are eliminated because the pronouns identified by the system are only personal pronouns.

- Noun phrases come before pronouns.

- Especially for pronouns in the form of clitic, the noun phrase attached by the clitic is not considered as a candidate.

  > [Putri Stéphanie] adalah anak bungsu Grace Kelly dan Rainier III dari Monako. Sesekali [ia] menjadi penyanyi, desainer pakaian renang, model, dan pemain sirkus.
  >
  > "[Princess Stephanie] is the youngest child of Grace Kelly and Rainier III from Monaco. Sometimes [she] becomes a singer, swimsuit designer, model, and circus player."

4 Data Annotation

Since there is no publicly available data of coreference resolution for Indonesian, we construct new data set in this research. We collect the data from Wikipedia in Indonesian language. We filter the Wikipedia pages that fulfill three criteria

1. The pages contain many noun phrases. We hypothesizes they are the Wikipedia pages discussing one of following topics. (i) fictional plots, e.g., subtitles for films, TV show episodes, and novel stories; (ii) biographies (incl. fictional characters); and (iii) historical events or important events.

2. The pages contain significant variation of pronoun and named-entity. We count the number of first, second, third person pronouns, and clitic pronouns in the document by applying string matching.We examine the number of named-entity using the Stanford CoreNLP NER Tagger (Manning et al., 2014) with a model trained from the Indonesian corpus taken from Alfina et al. (2016).

3. The Wikipedia texts have length of 500 to 2000 words.

Type of Mention	Frequency
Named-entity	6934
Pronoun	5736
Other noun phrase	3410
N/A	380

Table 1: Statistics of Mention Types in Documents

We sample 201 of pages from subset of filtered Wikipedia pages. We hire five annotators who are undergraduate student in Linguistics department. They are native in Indonesian. Annotation is carried out using the Script d'Annotation des Chanes de Rfrence (SACR), a web-based Coreference resolution annotation tool developed by Oberle (2018). From the 201 texts, there are 16,460 mentions tagged by the annotators. The distribution of mentions can be seen in Table 1

5 Experiments

5.1 Evaluation Metric

The multi-pass sieve coreference resolution system is evaluated on annotated data using two evaluation metrics commonly used for Coreference resolution, namely MUC and BCUBED F-measure.

MUC Evaluation Metrics The MUC-Link Based F-Measure metric was introduced in Message Understanding Coreference 6, and has become the standard for one of the most commonly used metrics in evaluating coreference resolution systems (Vilain et al., 1995). This metric considers the correlation relationship as a chain, where each mention is connected to a maximum of two chains. Basically, the MUC metric calculates the number of partitions required for the chain in the result-set to match the chain in the gold-standard.

$$partition(cluster, Set) = \{e | e \in cluster \& \exists s \in Set : e \in s\}$$

Precision, recall and F1-score for the MUC metric is calculated by the following equation:

$$Precision(G, R) = \sum_{r \in R} \frac{|r| - |partition(r, G)|}{|r| - 1}$$

$$Recall(G, R) = \sum_{g \in G} \frac{|g| - |partition(g, R)|}{|g| - 1}$$

$$F1 = \frac{2 * Precision(R, G) * Recall(G, R)}{Precision(R, G) + Recall(G, R)}$$

The weakness of MUC F-measure is that it is non-discriminatory. This is due to the absence of weight differences in link errors when a system over-merges a document (Luo, 2005).

BCUBED Evaluation Metrics The BCUBED is a mention-based metric (Bagga and Baldwin, 1998). BCUBED evaluation is done by calculating precision and recall for each mention M, then calculating the final result by weighted-sum of each precision and recall.

$$Precision_M = \frac{|M_{Result} \cap M_{Gold}|}{|M_{Result}|}$$

$$Recall_M = \frac{|M_{Result} \cap M_{Gold}|}{|M_{Gold}|}$$

$$FinalPrecision = \sum_{i=1}^{N} w_i * Precision_i$$

$$FinalRecall = \sum_{i=1}^{N} w_i * Recall_i$$

Weight is generally defined as $1/N$, where N is the number of noun phrases in the document. The F1-score calculation for the BCUBED metric uses a formula that is fundamentally the same as the calculation for the MUC metric as follows:

$$F1 = \frac{2 * FinalPrecision * FinalRecall}{FinalPrecision + FinalRecall}$$

5.2 Result and Analysis

Table 2 shows that, in general, there is an increasing trend of F1-score on the MUC metric along with the number of passes used. This is due to an increase in recall when one sieve is added one by one. However, when the system gets the best MUC F-measure value of 72.74% if all passes are implemented, the best BCUBED F-measure value is 52.18% when only the first three passes are implemented.

Based on Table 2, coreference resolution has the best precision in the implementation of Pass 1 and 2. This is different from the evaluation results of the Multi-Pass Sieve model in English and the Multi-Pass Sieve concept which depicts that the highest precision is at the top layer. This is due to cases in article text where nominal phrases with the same string refer to different entities. A common case of this allegation is generally the use of the last name to refer to a person, where people with the same surname will be considered the same entity.

Viewed from the group of features per sieve, string similarity feature is a feature that contributes the highest with a recall increase of 21.72% for MUC and 30.03% for BCUBED. The pronoun resolution is also a major contributing layer with recall increases of 33.08% for MUC and 21.99% for BCUBED, despite the decrease in precision and F1 in the BCUBED metric.

The difference in the F-measure trend of the MUC and BCUBED metrics is caused by the dif-

Pass	MUC			BCUBED		
	P	**R**	**F1**	**P**	**R**	**F1**
1	67.70	21.72	31.91	**94.11**	30.03	45.53
1,2	69.18	30.65	41.40	91.53	34.95	49.11
1,2,3	67.60	40.90	49.98	77.58	42.77	**52.18**
1,2,3,4	67.30	41.73	50.52	75.42	43.55	52.07
1,2,3,4,5	67.25	41.75	50.52	75.31	43.56	52.03
1,2,3,4,5,6	**71.36**	**74.83**	**72.74**	46.69	**65.55**	50.00

Table 2: Multi-Pass Sieve evaluation with MUC and BCUBED metrics

ference in the way the two metrics evaluate two clusters that are merged into one (over-merging). The BCUBED metric evaluates based on mentions, thus penalizing the cases of the merged cluster. The MUC metric evaluates based on mention-links, so that the number of partitions penalized for the two merged clusters is only one link, causing a fairly small penalty.

The decrease in precision in the BCUBED metric is due to the system's inability to fully detect the antecedents of plural pronouns. Increased precision can be done by increasing the compatibility of attributes between mentions, such as the number and gender attributes which are quite difficult to identify in Indonesian. On the other hand, the head match variation feature, namely proper head match and relaxed head match, does not have much effect on system performance, because the core noun attributes are similar to the previous sieve (strict head match) but cannot take advantage of additional rules such as attribute agreement.

6 Conclusion

We conclude that the Multi-Pass Sieve approach provides a strong baseline performance for coreference resolution task in Indonesian language. The Multi-Pass Sieve approach achieves MUC F-measure up to 72.74% and BCUBED F-measure up to 52.18%. The use of the exact string match feature provides high precision for the Multi-Pass Sieve model, while the increase in recall is also influenced by the Pronoun Sieve. The highest F-measure results for the MUC metric were obtained by Multi-Pass Sieve for the combined implementation of Pass 1-6 and Pass 1-3 models for the BCUBED metric. On the other hand, the proper head match and relaxed head match features do not appear to have a major impact on system performance, contributing to a performance increase of around 1%.

There is still a lot of room to improve the coreference resolution system for Indonesian. In the future, we plan to implement an end-to-end model coreference resolution system, in which the resolved phrases or words can be detected automatically using a more perfect chunking and named-entity recognition system.

References

Ika Alfina, Ruli Manurung, and Mohamad Ivan Fanany. 2016. Dbpedia entities expansion in automatically building dataset for indonesian ner. In *2016 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, pages 335–340.

Hasan Alwi, Soenjono Dardjowidjojo, Hans Lapoliwa, and Anton M. Moeliono. 2010. *Tata Bahasa Baku Bahasa Indonesia*, 8th edition. PT Balai Pustaka (Persero), Jakarta, Indonesia.

Amit Bagga and Breck Baldwin. 1998. Algorithms for scoring coreference chains. In *In The First International Conference on Language Resources and Evaluation Workshop on Linguistics Coreference*, pages 563–566.

Indra Budi, Stephane Bressan, and Nasrullah. 2006. Co-reference resolution for the indonesian language using association rules. In *iiWAS*.

Chen Chen and Vincent Ng. 2012. Combining the best of two worlds: A hybrid approach to multilingual coreference resolution. In *Joint Conference on EMNLP and CoNLL - Shared Task*, pages 56–63, Jeju Island, Korea. Association for Computational Linguistics.

Lynette Hirschman and Nancy Chinchor. 1998. Appendix F: MUC-7 coreference task definition (version 3.0). In *Seventh Message Understanding Conference (MUC-7): Proceedings of a Conference Held in Fairfax, Virginia, April 29 - May 1, 1998*.

Septina Dian Larasati. 2012. Handling Indonesian clitics: A dataset comparison for an Indonesian-English statistical machine translation system. In *Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation*, pages 146–152, Bali, Indonesia. Faculty of Computer Science, Universitas Indonesia.

Heeyoung Lee, Yves Peirsman, Angel Chang, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2011. Stanford's multi-pass sieve coreference resolution system at the CoNLL-2011 shared task. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, pages 28–34, Portland, Oregon, USA. Association for Computational Linguistics.

Xiaoqiang Luo. 2005. On coreference resolution performance metrics. In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pages 25–32, Vancouver, British Columbia, Canada. Association for Computational Linguistics.

Christopher Manning, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky. 2014. The Stanford CoreNLP natural language processing toolkit. In *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 55–60, Baltimore, Maryland. Association for Computational Linguistics.

Bruno Oberle. 2018. SACR: A drag-and-drop based tool for coreference annotation. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Cheon-Eum Park, Kyoung-Ho Choi, and Changki Lee. 2014. Korean coreference resolution using the multi-pass sieve. *Journal of KIISE*, 41:992–1005.

Desmond Darma Putra, Abdul Arfan, and Ruli Manurung. 2008. Building an indonesian wordnet. In *Proceedings of the 2nd International MALINDO Workshop*.

Karthik Raghunathan, Heeyoung Lee, Sudarshan Rangarajan, Nathanael Chambers, Mihai Surdeanu, Dan Jurafsky, and Christopher Manning. 2010. A multi-pass sieve for coreference resolution. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 492–501, Cambridge, MA. Association for Computational Linguistics.

Ander Soraluze, Olatz Arregi, Xabier Arregi, and Arantza Díaz de Ilarraza. 2019. Euskor: End-to-end coreference resolution system for basque. *PLOS ONE*, 14(9):1–25.

Totok Suhardijanto, Rahmad Mahendra, Zahroh Nuriah, and Adi Budiwiyanto. 2020. The framework of multiword expression in Indonesian language. In *Proceedings of the 34th Pacific Asia Conference on Language, Information and Computation*, pages 582–588, Hanoi, Vietnam. Association for Computational Linguistics.

Gilang Julian Suherik and Ayu Purwarianti. 2017. Experiments on coreference resolution for indonesian language with lexical and shallow syntactic features. In *2017 5th International Conference on Information and Communication Technology (ICoIC7)*, pages 1–5.

Rhea Sukthanker, Soujanya Poria, Erik Cambria, and Ramkumar Thirunavukarasu. 2020. Anaphora and coreference resolution: A review. *Information Fusion*, 59:139–162.

Marc Vilain, John Burger, John Aberdeen, Dennis Connolly, and Lynette Hirschman. 1995. A model-theoretic coreference scoring scheme. In *Proceedings of the 6th Conference on Message Understanding*, MUC6 '95, page 45–52, USA. Association for Computational Linguistics.