A Logical Approach to Sentiment Analysis

Niklas Christoffer Petersen

Kongens Lyngby 2012
IMM-MSc-2012-????

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk

IMM-MSc-2012-????

Summary (English)

The goal of the thesis is to ...

Summary (Danish)

Målet for denne afhandling er at ...

Preface

This thesis was prepared at the department of Informatics and Mathematical Modelling at the Technical University of Denmark in fulfilment of the requirements for acquiring an MSc in Computer Science and Engineering.

The thesis deals with ...

The thesis consists of ...

Kgs. Lyngby, September 30, 2012

Niklas Christoffer Petersen

Acknowledgements

I would like to thank my....

Contents

Chapter 1

Introduction

In the UK, online retail sales account for over 10% of purchases, and their growth rate is markedly outstripping store-based sales [13]. Many customers express their opinions about products through online reviews. These reviews are key for marketing intelligence since they contain valuable information. Popular products may have hundreds of reviews making it hard for the customer to find the information they require, and as a result there is a need to automatically classify this data. This can benefit both customers and manufacturers of products. Customers can see what other consumers thought about the products', viewing the products strengths and weaknesses. Manufacturers can then see where their product falls short in order to improve it, and also they can compare their products to other competitive products

Cite, rewrite, *Sentiment Analysis of Customer Reviews ...*

Noisy unstructured text data is generated in informal settings such as online chat, emails, blogs, customer feedbacks and reviews. These texts have the potential to act as rich sources for raw inputs to market research and knowledge discovery. Since Internet is a crucial driving force in today's world, these texts are rich pointers to the collective opinion of the global population on almost every topic.

Cite, rewrite, *Opinion Mining From Noisy Text Data*

Feedback, in form of product and service reviews, can thus be heigly valuable information for companies. Also political opinions are of high value for both governments and their the oppositions. The interest in such opinions is far from recent, and is a well etablished subfield of the *psychometrics* and has strong scientific grounds in both psychology and statistics.

Introduce *sentiment analysis* (also sometimes refered as *opinion mining*)

1.1 Classical data collection

One of the most used approaches to collect data for opinion analyses is through questionnaire surveys. Most of us are familiar with such surveys, where the subject is forced to answer questions with a fixed scale. For instance, given the statement "The rooms at the Swissôtel Hotel are of high quality.", a subject must answer by selecting one of a predefined set of answers, e.g. as shown in figure 1.1

1. Strongly disagree
2. Disagree
3. Neither agree nor disagree
4. Agree
5. Strongly agree

Figure 1.1: Likert scale.

Such scales, where the subject indicates the *level of agreement*, are know as *Likert scales*, originally presented by Likert [1932], and has been one of the favourite methods of collection data for opinion analyses cf. [**?**]. Other scales are also widely used, for instance the *Guttman scale* [**?**], where the questions are binary (yes/no) and ordered such that answering yes to a questions implies the answer yes to all questions ordered below this. Thus the answer on a Guttman scale can be captured by a single index. An example of an Guttman scale is shown in figure 1.2.

1. I like eating out
2. I like going to restaurants
3. I like going to themed restaurants
4. I like going to Chinese restaurants
5. I like going to Beijing-style Chinese restaurants

Figure 1.2: Guttman scale.

Given a set of answers, the result of such surveys are fairly easy to compute. At its simplest it can be an average of the answers, however mostly it is also interesting to connect the questions – for instance how does subjects' answer to the above statement influent their answer to the statement "The food at the Swissôtel Restaurant are of high quality.", etc.

List some shortcommings of this approach, e.g. predefined frame for feeedback etc., it is often hard to get people to participate, etc.

1.2 Natural language data collection

In this thesis it is argued that a far more natural way for subjects to express their opinions is through their most natural communication form, i.e. their language, eigther in written or spoken form.

The initiative for such data collection could be *opinion seeking queries* as the one shown in (1.1). Such queries are intended to ensure succinct reviews that clearly relate to the *entity* in question (e.g. product or service) with respect to a specific *topic of interest.* Clearly the queries could be formulated in a more friendly and *call to action* manner, as shown in (1.2)

Holiday Inn, London: price (1.1)

What do you think about pricing at the Holiday Inn, London? (1.2)

This method might not seem that different from that of the previously mentioned Likert scales, but it still allows the reviewer to answer with a much broader sentiment and lets the reviewer argue for his/hers answer as shown in the examples (1.3-1.4).

The price is moderate for the service and the location. (1.3)

Overall an above average hotel based on location and price but not one for a romantic getaway! (1.4)

More passive sources could also be considered, including posts on social networking services and microblogging services (e.g. Facebook[1] and Twitter[2]). This though introduces the need for efficient candidate filtering, as the posts in general of cause are not constrained to a specific entity or topic of interest. However it also significantly increases the data quantity, which in turn can yield a more precise analysis. Since the author of the post might never realize that the post is being used for the purpose of opinion analysis it also raises ethical issues. Larger texts, such as blot posts, could also be considered, however the contextual aspects of large, contiguous texts often makes interpretation extramly complex, thus making it a most difficult task to extract opinions on a specific entity.

It is also worth mentioning the possibility to collect answers such as (1.3, 1.4) in spoken form, which would give perhaps the most natural interaction. However for the purpose of this thesis it is proposed to solely focus on language in written form, as spoken form introduces a lot of complexity due to the speech recognition needed, for instance efficient audio sampling and analysis, speaker dependence (e.g. dialects), etc. Further more the interpretation of spoken language is also highly complex, since the emotional perception of the speaker must be considered in order to detect for instance ironic statements.

The overall goal is thus to perform a *sentiment analysis* of a set of small, succint, review texts with respect to a subject, and yield an normalized score.

[1]Facebook, `http://www.facebook.com/`
[2]Twitter, `http://www.twitter.com/`

Fit

We will present the problematics that arises when trying to extract the semantic of texts from a language with a large vocabulary, and present several computational approaches for solving (at least partly) these issues.

1.3 Related work

In the following notable related work on sentiment analysis are briefly presented, and based on this it is argued that there are two main approaches for sentiment analysis of written texts, namely using *formal systems* and respecitively using *statistical methods*.

- **Formal approaches** With this approach the language of the texts to analyse is modeled as a formal language, i.e. using a formal grammar. This allows a syntactical analysis of the texts, yielding the structures of the texts, e.g. sentences, phrases, and words. Semantic information is then extractable by augmenting and inspecting these structures.

  The result of the semantic analysis is then subject to the actual sentiment analysis ...

  Cites

- **Statistical approaches** With a statistical approach a probabilistic model is constructed, and trained against a training data set, configuring the perameters of the model (of which there can be an tremendous amount of). The model is then applied to the actual data set of which an analysis is desired.

  can extract features from the input text. These features can then be analyzed in order to extract semantic properties.

  Cites: Language models, positional language model, HHM

Both of these has been considered as the foundation of the analysis of the text reviews, however it is proposed to follow a logic approach for several reasons:

- Given that it is a formal system, it can be modeled closely and hopefully also efficiently in the *proof of concept* implementation.
- It is questionable whether the entropy of each topic in the chosen dataset actually allows feature extraction on a significant level. Given a set of text on a topic is divided into a training set and a test set, it is doubtful that the trained model would be able to capture features from the test set, since the topics only contains approximately a hundred texts.
- The undersigned has far more experience in the field of logic systems, than in the field of statistics, thus a higher level

Notice that even though it is proposed to follow a logic approach, it is still the intention to retain the proposed solutions for misspellings and minor grammatical errors, which utilize some probabilistic properties, e.g. *match scores*, etc. However these should be seen solely as preprocessing steps needed, in case the input text cannot be analyzed directly, e.g. as a failover precaution. Thus for perfect texts, the analysis will be purely logic.

1.4 Using real data sets

For the presented solution to be truely convincing we wish to present a *proof of concept* implementation that shows at least some of the desired capabilities. However, for such product to be demonstrated properly, real data is required. Testing in on some tiny pseudo data set constructed for the sole purpose of this demonstration would not be convincing.

Move to chap. 2?

Dealing with major grammatical errors, such as wrong word order is a much harder problem, since even small changes in for instance the relative order of subject, object, verb etc. may result in an major change in interpretation. Thus it is proposed, only to focus on minor grammatical errors such as incorrect form. Such corrections could be made achievable by linking different forms of the same word.

> Blog texts on the other hand suffer from ill-composed sentences, arbitrary punctuations or insertions of characters, irrational capitalization etc. Spelling correction, abbreviation and case restoration mechanisms for noisy text have been addressed in [3, 4]. We have employed variations of these techniques for our work.

Should be 3-5 pages: Motivation, Project goals, and Thesis structure

Chapter 2

Logical sentiment analysis

For the purpose of this thesis a *sentiment analysis* is defined as follows:

Definition 2.1 a sentiment analysis $\mathcal{A}$ is a computation on a review text $T \in \Sigma^\star$ with respect to a subject $s \in S$, where $\Sigma^\star$ denotes the set of all permissible texts in the language. The result is an normalized score as shown in (2.1). The yielded score should reflect the *polarity* of the subject in the text, i.e. whether the overall opinion is positive (a score close to ι), negative (a score close to $-\iota$), or neutral (a score close to 0).

$$\mathcal{A} : \Sigma^\star \to S \to [-\iota; \iota] \tag{2.1}$$

It should be evident that this computation is far from triviel, and constitutes the cornerstone of this thesis. There are several steps needed, if such computation should yield any reasonable result. As mentioned in the introduction the goal is a logical approach for achieving this. The following outlines the overall steps to be completed, their associated problematics in this process, and succintively presents different approaches to solve each step. The chosen approach for each step will be presented in much more details in later chapters. Finally ...

Something about data acquisition of test-data

2.1 Syntactic analysis

The first step is to determine the grammatical structure of the input texts with respect to the rules of the English language. It is expected that the reader is familiar with English grammar rules and syntactic categories, including phrasal categories and lexical categories (also called parts of speech). As mentioned erlier it is essential that the chosen solution is able to cope with *real* data, collected from actual review scenarios. This implies a robust syntactic analysis accepting a large vocabulary and a wide range of sentence structures. In order to calculate the actual polarity it is essential to have semantic annotations on the lexical units. It is argued that a feasible and suitable solution is to use a grammar that is *lexicalized*, i.e. where the rules are essentially langauge independent, and the syntactic properties are derived from a lexicon. Thus the development of a lexicalized grammar is mainly a task of acquiring a sutable lexicon for the desired language.

Even though the task of syntactic analysis now is largely reduced to a task of lexicon acquisition, which will be addressed in a moment, there are still general concerns that are worth acknowledging. Simply identifying the different sentences and lexical units in a text can yield a difficult task. Consider for intance the text (2.2) which is taken from the Wall Street Journal corpus [Paul and Baker, 1992]. There are six periods in it, but only two of them indicates sentence bounderies, and delimits the passage into its two sentences.

Pierre Vinken, 61 years old, will join the board as a nonexecutive director Nov. 29. Mr. Vinken is chairman of Elsevier N.V., the Dutch publishing group. (2.2)

Hockenmaier *et al.* [2004, p. 108-110] identifies several issues in being able to efficiently handle natural language texts solely with lexicalized grammars, mainly due to the need for entries for various combinations of proper nouns, abbreviated terms, dates, numbers, etc. Instead they suggest to use pattern matching and statistical techniques as a preprocessing step, for which efficient components exists, which translate into reduced complexity for the actual syntactic analysis.

Fortunately the domain of small review texts allows some restrictions and assumptions, that at least will ease some of the issues in this context. For instance it seems like a valid assumption that most of the review texts will consists of at most two sentences. Its is argued that this indeed is achievable by sufficient intructing and constraining the reviewers duing data collection, e.g. only allowing up to a curtain number of characters. It is also argued that the use of proper nouns and specific dates can be fairly limited, in that a context have already been established for the reviewer cf. section 1.2. However the domain of small review texts also introduce concerns that are absent from other domains, including the possebility of incorrect grammar and spelling, since the texts comes unedited from humans with varying English skills. A solution that would only work on *perfect texts* (i.e. texts of sentences with completely correct grammar and spelling) would not be adequate. In order to at least try to handle minor misspellings it is intended to use algoritms that can select alternatives from the lexicon. Reasons for this could be that word is simply unpresent from the system's vocabulary (e.g. misspelled), or on a grammatical incorrect form (e.g. wrong person, gender, tense, case, etc.).

2.1.1 Mildly context-sensitive grammars

There exists formal proofs that some natural language structures requires formal power beyond *context-free grammars* (CFG), i.e. [Shieber, 1985] and [Bresnan *et al.*, 1982]. Thus the search for grammars with more expressive power has long been a major study within the field of computational linguistics. The goal is a grammar that is so restrive as possible, allowing efficient syntactic analysis, but still capable of capturing these structures. The class of *mildly context-sensitive grammars* are conjectured to be powerful enough to model natural languages while remaining efficient with respect to syntactic analysis cf. [Joshi *et al.*, 1990].

Different grammar formalisms from this class has been considered, including *Tree Adjunct Grammar* (TAG) [Joshi *et al.*, 1975], in its lexicalized form (LTAG), *Head Grammar* (HG) [Pollard, 1984] and *Categorial Grammar* (CG) [Steedman, 1998]. It has been shown that these are all equal in expressive power by Vijay-Shanker and Weir [1994]. The grammar formilism chosen for the purpose of this thesis is *Combinatory Categorial Grammar* (CCG), pioneered largely by Steedman [2000]. CCG adds a layer of combinatory logic onto pure CG.

Find original 1985/1988 article

Explain the choice of CCG

Chapter 3 will formally introduce the CCG in much more detail.

2.1.2 Lexicon acquisition

As mentioned earlier, acquiring a sutable lexicon is crucial in the development of a lexicalized grammar, especially when the grammar is to accept a large vocabulary and a wide range of sentence structures.

There exists several wide coverage CCG lexicons, most notable *CCGbank*, compiled by Hockenmaier and Steedman [2007] by translating almost the entire Penn Treebank [Marcus *et al.*, 1993], which contains over 4.5 million words. The result is a higly covering lexicon, with some entries having assigned over 100 different lexical categories. Unfortunately these lexicons are not free to use, and it has not been possible to fund a license.

2.2 Semantic analysis

2.2.1 Using real data-sets

Continue...

2.3 Data acquisition

In order to sucesfully perform the proposed computation, and thus the sentiment analysis

2.4 Tagged corpora / Part-of-speech tagging

Since English is

2.4.1 The Brown Corpus

The Brown Corpus was compiled by Francis and Kucera [1979] by collecting written works printed in United States during the year 1961. The corpus consists of just over one million words taken from 500 American English sample texts, with the intension of covering a highly representative variety of writing styles and sentence structures.

Notable drawbacks of the Brown Corpus include its age, i.e. there are evidently review topics where essential and recurring words used in present day writing was not coined yet or rarely used back 50 years ago. For instance does the Brown Corpus not recognize the words *internet*, *hotspot*

sentences will containing words has found it's way into comon that

Other corpora has been considered

> *As with English around the world, the English language as used in the United Kingdom and the Republic of Ireland is governed by convention rather than formal code: there is no equivalent body to the Académie française or the Real Academia Española, and the authoritative dictionaries (for example, Oxford English Dictionary, Longman Dictionary of Contemporary English, Chambers Dictionary, Collins Dictionary) record usage rather than prescribe it. In addition, vocabulary and usage change with time; words are freely borrowed from other languages and other strains of English, and neologisms are frequent.*
>
> *http://en.wikipedia.org/wiki/British_English#Standardisation*

2.5 Test dataset

2.5.1 The Opinosis Dataset

The suggested data set to use is the *Opinosis Dataset*, originally used by [Ganesan *et al.*, 2010]. The data set consists of texts from actual user reviews on a total of 51 different topics. The topics are ranging over different objects, from consumer electronics (e.g. GPS navigation, music players, etc.) to hotels, restaurants and cars. For most of the objects, reviews are covered by multiple topics. For instance a specific car is covered by the topics *comfort*, *interior*, *mileage*, *performance*, and *seats*.

It has been hard to find any real alternatives for the *Opinosis Dataset* for several reasons: Most collected reviews are commercial, and thus not free to use; furthermore the *Opinosis Dataset* also contains summerized texts for each of its topics, which are constructed by manual, human interpretation. The latter allow a straight approach for comparison of any results the proposed system will yield.

The hotel buffet had fabulous food. (2.3)

Very friendly servers and nice selection of food at a reasonable price. (2.4)

Room service was extortionate though, very very expensive,
so we didnt bother, as food outlets a few minutes walk away. (2.5)

The texts (2.3) to (2.5) show actual extracts from the data set for a topic on food quality on the Swissôtel Restaurant. While (2.3) is a valid declarative sentence, (2.4) is not, since it lacks a subject (i.e. the restaurant). A coarse review of the text in the dataset reveals that missing subjects are a repeating issue. This might not seem that odd, since many people would implicitly imply the subject from the topic that they are reviewing. Thus text missing subjects can in many cases still be considered as valid sentences with minimal effort. The text (2.5) is on the other hand missing a transitive verb (presumably *are*) from the subordinate clause. In cases where such savere gramatically errors occurs it is sugested to ignore the clause, and try only to analyse the main clause. Furthermore the text (2.5) use repeated adverbs (e.g. *very very*) to express intensification, however it should not be any major concern that a verb or adjective are modified multiple times by the *same* adverb, but the intended intensification will probably not be included in the semantic analysis. Thus formalizing such a grammer is mostly a tak od designing such lexicon.

As evendent from these examples far from all texts in the dataset are valid sentences.

Chapter 3

Combinatory categorial grammar

We expect such an algorithm to calculate a match score, that is a weighted average over several metrics. Given below are methods for calculating scores for some evident metrics.

- Symbolic similarity – at its most basic form we can consider a sample string (i.e. a word from an input text) against the system's vocabulary using approximate string matching algorithms such as the *Levenshtein distance* as described by [Wagner and Fischer, 1974].

- Pronunciation similarity – it is an valid assumption that many misspellings still share a majority of the pronunciation with the intended word, i.e. they are approximately homophone. Thus comparing the phonetic properties of an sample string with possible matches can in cases correct misspellings. The *Soundex algorithm* by Robert C. Russell and Margaret K. Odell, as described by [Knuth, 1998, p. 391–92], is a simple, yet power full approach for this purpose.

3.1 Features and agreement

In order to ensure that the parsed phrases indeed follows correct English grammar, it is not enough to only consider the phrase structure with respect to the word classes. It is also necessary to

As stated it is essential to categorize phrases, however it is important to notice, that each category can have numerous arguments i.a. denoting features that apply. For instance Bob Carpenter [1995] states features for person (e.g. 1st, 2nd, or 3rd), number (e.g. singular or plural), and case (e.g. subject or object). The set of features that may apply is language dependent, for instance most indo-european langauges1 has gender features (English being an exception2), but while Danish use common and neuter classes, other languages like German use masculine, feminine, and neuter classes. The important is thus solely that the set of features is always finite – limited by the specific langauge.

[van Eijck and Unger, 2010]

3.2 Find a good title

The initial attempt is simply to construct a parser that

> There are three basic ways to build a shift-reduce parser. Full LR(1) (the 'L' is the direction in which the input is scanned, the 'R' is the way in which the parse is built, and the '1' is the number of tokens of lookahead) generates a parser with many states, and is therefore large and slow. SLR(1) (simple LR(1)) is a cut-down version of LR(1) which generates parsers with roughly one-tenth as many states, but lacks the power to parse many grammars (it finds conflicts in grammars which have none under LR(1)).
>
> LALR(1) (look-ahead LR(1)), the method used by Happy and yacc, is tradeoff between the two. An LALR(1) parser has the same number of states as an SLR(1) parser, but it uses a more complex method to calculate the lookahead tokens that are valid at each point, and resolves many of the conflicts that SLR(1) finds. However, there may still be conflicts in an LALR(1) parser that wouldn't be there with full LR(1).

The state S_τ ...

Formally a rule $\mathcal{R}_\tau$, for the state type τ, is a transformation from a state $s \in \mathcal{S}_\tau$ onto a new set of states $\mathcal{S}'_\tau \subset \mathcal{S}_\tau$ cf. 3.1.

$$\mathcal{R}_\tau : \mathcal{S}_\tau \to \mathcal{P}(\mathcal{S}_\tau) \tag{3.1}$$

The state type for analysing CCGs is a 2-tuple, where P is a totally ordered set of ...,

$$\mathcal{S}_{\mathrm{CCG}} : \mathcal{P}(T) \times \mathcal{P}(\mathcal{P}(T))$$

$$\mathcal{R}_{\mathrm{CCG}}^{\mathrm{shift}} \tag{3.2}$$

If all rules in the set is monotone, then the parsing will terminate

Chapter 4

Implementation

It was chosen to use the functional programming language *Haskell* for implementing the *proof of concept* program. In the following sections key aspects of the implementation will be presented. For the complete source code for the implementation please see appendix **??**.

The reason Haskell, specifically the *Glasgow Haskell Compiler*, was chosen as programming language and platform, was the ability to ...

4.1 Tokenizer and tagger

The tokenizer has a very simple task, namely to convert an input string to a list of tokens (lower case words) that represent the symbols of the language. An example of the transformation is shown in (4.1).

$$\text{“Put the pyramid onto the table.”} \rightarrow [\mathbf{put}, \mathbf{the}, \mathbf{pyramid}, \mathbf{onto}, \mathbf{the}, \mathbf{table}] \quad (4.1)$$

Appendix A

Stuff

This appendix is full of stuff ...

Bibliography

[Bresnan *et al.*, 1982] J. Bresnan, R. M. Kaplan, S. Peters, and A. Zaenen. Cross-Serial Dependencies in Dutch. *Linguistic Inquiry*, 13(fall):613–635+, 1982.

[Francis and Kucera, 1979] W. Nelson Francis and Henry Kucera. Brown corpus manual. Technical report, Department of Linguistics, Brown University, Providence, Rhode Island, US, 1979.

[Ganesan *et al.*, 2010] Kavita Ganesan, Cheng Xiang Zhai, and Jiawei Han. Opinosis: A graph based approach to abstractive summarization of highly redundant opinions. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING '10)*, 2010.

[Hockenmaier and Steedman, 2007] Julia Hockenmaier and Mark Steedman. CCGbank: A Corpus of CCG Derivations and Dependency Structures Extracted from the Penn Treebank. *Computational Linguistics*, 33(3):355–396, 2007.

[Hockenmaier *et al.*, 2004] Julia Hockenmaier, Gann Bierner, and Jason Baldridge. Extending the coverage of a ccg system. *Journal of Language and Computation*, 2:165–208, 2004.

[Joshi *et al.*, 1975] Aravind K. Joshi, Leon S. Levy, and Masako Takahashi. Tree adjunct grammars. *Journal of Computer and System Sciences*, 10(1):136–163, 1975.

[Joshi *et al.*, 1990] Aravind K. Joshi, K. Vijay Shanker, and David Weir. *The Convergence Of Mildly Context-Sensitive Grammar Formalisms*, 1990.

[Knuth, 1998] Donald E. Knuth. *The Art of Computer Programming, vol. 3: Sorting and Searching*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 2nd edition, 1998.

[Likert, 1932] Rensis Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 22(140):1–55, 1932.

[Marcus *et al.*, 1993] Mitchell P. Marcus, Mary Ann Marcinkiewicz, and Beatrice Santorini. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics*, 19(2):313–330, 1993.

[Paul and Baker, 1992] Douglas B. Paul and Janet M. Baker. The design for the Wall Street Journal-based CSR corpus. In *Proceedings of the workshop on Speech and Natural Language*, pages 357–362, Stroudsburg, PA, USA, 1992. Association for Computational Linguistics.

[Pollard, 1984] Carl Pollard. PhD thesis, 1984.

[Shieber, 1985] Stuart M. Shieber. Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8(3):333–343, 1985.

[Steedman, 1998] Mark Steedman. *Categorial Grammar*, 1998.

[Steedman, 2000] Mark Steedman. *The Syntactic Process*. The MIT Press, 2000.

[van Eijck and Unger, 2010] Jan van Eijck and Christina Unger. *Computational Semantics with Functional Programming*. Cambridge University Press, New York, NY, USA, 1st edition, 2010.

[Vijay-Shanker and Weir, 1994] K. Vijay-Shanker and David J. Weir. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory*, 27:27–511, 1994.

[Wagner and Fischer, 1974] Robert A. Wagner and Michael J. Fischer. The string-to-string correction problem. *J. ACM*, 21(1):168–173, January 1974.